


9-2016

NeuroEthics: NeuroLaw

Stephen J. Morse

University of Pennsylvania Law School

Follow this and additional works at: http://scholarship.law.upenn.edu/faculty_scholarship

 Part of the [Bioethics and Medical Ethics Commons](#), [Biological Psychology Commons](#), [Criminal Law Commons](#), [Ethics and Political Philosophy Commons](#), [Genetics Commons](#), [Health Law and Policy Commons](#), [Law and Philosophy Commons](#), [Law and Psychology Commons](#), [Mental Disorders Commons](#), [Other Neuroscience and Neurobiology Commons](#), and the [Public Law and Legal Theory Commons](#)

Recommended Citation

Morse, Stephen J., "NeuroEthics: NeuroLaw" (2016). *Faculty Scholarship*. 1726.
http://scholarship.law.upenn.edu/faculty_scholarship/1726

This Article is brought to you for free and open access by Penn Law: Legal Scholarship Repository. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of Penn Law: Legal Scholarship Repository. For more information, please contact PennlawIR@law.upenn.edu.

NeuroEthics: NeuroLaw
Stephen J. Morse
Oxford Handbook Online (Philosophy)

Abstract

This entry discusses whether the findings of the new neuroscience based largely on functional brain imaging raise new normative questions and entail normative conclusions for ethical and legal theory and practice. After reviewing the source of optimism about neuroscientific contributions and the current scientific status of neuroscience, it addresses a radical challenge neuroscience allegedly presents: whether neuroscience proves persons do not have agency. It then considers a series of discrete topics in neuroethics and neurolaw, including the “problem” of responsibility, enhancement of normal functioning, threats to civil liberty, competence, informed consent, end of life issues, and the ethics of caution. It suggests that the ethical and legal resources to respond to the findings of neuroscience already exist and will do so for the foreseeable future.

Keywords

neuroscience, neuroethics, neurolaw, agency, responsibility, enhancement, civil liberty, competence, informed consent

I. Introduction

Neuroethics and neurolaw, which have become subjects of intense attention in the recent past (e.g., Chatterjee and Farah, 2013; Farah, 2010; Garland, 2004; Glannon, 2007, 2013; Goodenough & Tucker, 2010; Hoffman, 2014; Illes, 2005; Illes and Sahakian, 2011; Jones, Schall & Shen, 2014; Jones & Ginther, 2015; Levy, 2007; Morse, 2004; Morse & Roskies, 2013; Pardo and Patterson, 2013; Roskies, 2004, 2016; Satel & Lilienfeld, 2013; Vincent, 2013) address two distinct sets of questions. The first are largely empirical, asking about the neural correlates and sometimes neural causes of ethical and legal judgment and decision making. The second set of questions addresses how the findings of neuroscience should influence ethical and legal theory and practice. Some doubt the independent existence of neuroethics and neurolaw. Such critics believe either that these alleged fields are simply questions within pure neuroscience or bioethics or that the field should be expanded to include contributions from the allied disciplines of cognitive science and psychology. There is merit to both types of critique, but this entry

will proceed on the assumptions that neuroethics and neurolaw are sufficiently independent areas of inquiry to justify independent treatment and that allied sciences are part of the field because cognitive, affective and social neuroscience, the branches of neuroscience most relevant to ethics and law, are necessarily tied to allied disciplines such as cognitive science and psychology. This entry will return to why the allied sciences should be included below in Part III.

The findings of the first set of neuroethical studies seldom have normative consequences for ethical or legal theory or practice. Let us call this branch of these disciplines, empirical neuroethics and neurolaw. For example, a recent, sophisticated neuroscience paper used a technique that permitted inferences about the causal role that certain brain regions may have played in mental events (Buckholz et al., 2015). It found, *inter alia*, that dorsolateral prefrontal cortex (DLPFC) activity is associated with changes in decisions about punishment but not about blameworthiness. Finding a selective causal role for DLPFC in norm enforcement is fascinating, but the paper recognized that it was normatively inert. In a related recent study that employed a very creative methodology, Ginther et al (2016) were able to dissociate the brain regions associated with harm, mental state and integrative harm/mental state evaluations in making third party punishment decisions. Once again, this is fascinating research, but it does not entail any particular view of what punishments should be assigned to the offender in the experimental scenarios. No inference about the propriety of particular blameworthiness or punishment norms is possible from a finding about which brain region of interest (ROI) seems implicated. Nor is it surprising that different psychological processes should be associated with different brain activity. For a final example, the noted neurophilosopher Patricia Churchland's recent book, Braintrust: What Neuroscience Tells Us About Morality (2012), provided a neuroscientific account of how morality is biologically undergirded, but it does not purport to claim based on neuroscience what the principles and rules should be.

No single empirical finding is likely to have normative implications, but even when various findings converge on a particular result, it is still often unclear what follows for guiding ethics or law. Perhaps the best example is the question of adolescent criminal responsibility and how adolescent offenders (and perhaps young adults) should be treated by the criminal and juvenile justice systems. Widely-accepted and often replicated consistent findings in cognitive science, developmental psychology and neuroscience have found that adolescent rational capacities and brain maturation differ significantly on average from those of adults and that there are differences within the adolescent years (e.g., Scott et al, 2016). There is no more powerful

example of how various findings with genuine moral and legal relevance have converged. Nonetheless, it is not clear whether and in what way adolescent offenders should be treated differently from adults. If one has a moral and legal theory that suggests an outcome based on the facts, then it is the moral or legal theory that is playing an essential role in deciding what to do. Unless that theory is uncontroversial and the conclusion about the implications of the facts for that theory are equally uncontroversial, however, it is still not clear what follows from those facts (see Berker, 2009 and Kamm, 2009, for particularly trenchant, scientifically-informed analyses of the normative significance of neuroscientific findings).

It is of course no surprise that all human behavior should have in part a neuroscientific foundation. After all, without the types of brains we have, morality would be a pure abstraction in the mind of no one and guiding no behavior. If neuroscience (or any other science) is able to discover incontrovertible truths about the essence of human nature or the limits of our capacities, such information will certainly be relevant to normative analysis, but such truths are unlikely to *per se* entail normative conclusions about how we should live together.

The second set of questions is explicitly normative. Let us call this branch normative neuroethics and normative neurolaw. For example, in a widely-noticed chapter, neuroscientists Joshua Greene and Jonathan Cohen argued that the increasingly mechanistic understanding of the brain/mind that neuroscience is producing will convince us that we are all simply victims of neuronal circumstances, that no one is genuinely responsible, and that retributive justice should be abandoned in favor of a purely consequential prediction/prevention scheme for the social control of dangerous behavior (Greene & Cohen, 2006). Such work is almost entirely normative and shall be the focus of this on-line entry. Normative neuroethics/neurolaw may be said to have two distinct but sometimes overlapping sub-branches. The first involves claims that neuroscientific findings *per se* entail normative consequences. The Greene/Cohen claim is of this type. The second involves whether neuroscientific findings raise new ethical or legal issues or should affect principles, doctrines and practices that are already familiar.

Law and ethics, which include neurolaw and neuroethics, are two of the primary institutions humans have devised to guide our interpersonal lives. They share this primary function with many other institutions, including custom, etiquette, and social norms. Each gives us reasons to behave one way or another as we pursue our lives together. Although normative ethics and law share action-guiding and value-creating functions, law is also backed by the coercive power of the state, so it plays a central role in and applies to the

lives of all, including those who may disagree with particular laws. It still gives them at least instrumental reasons to conform. Although most theorists believe that morality and law are not co-extensive, the law often adopts rules that are primarily supported by moral considerations. Think of the core prohibitions of criminal law, which criminalize force, fraud and theft. Consequently, the most powerful practical role that normative neuroethics may play is to influence law-makers.

This entry begins by speculating about why, despite the limited achievements of behavioral (cognitive, social and affective) neuroscience to date, so many philosophers and lawyers are making what are apparently inflated claims about the implications of neuroscience for their fields (Morse, 2006, 2013). It next offers a brief overview of the methodology of behavioral neuroscience and the limitations on what we know at present and are likely to know in the future. In particular, it discusses whether neuroscience sheds new light on the relation between brain, mind and action. The entry then discusses a neuroscientific challenge to both normative disciplines, which is the claim that we are not agents at all. Both normative disciplines assume that ethics and law in large part are action-guiding. If we are not agents who can be guided by reason, what is the status of ethics and law? Perhaps the reasons ethics and law provide are epiphenomenal and thus they would have no genuinely causal action-guiding potential. This is a foundationally radical challenge that needs to be addressed before turning to normative neuroethics and neurolaw.

The entry then turns to normative neuroethics and neurolaw. The central thesis is that although the new neuroscience, especially fueled by non-invasive neuroimaging, is a new science using a new technology, at present and for the foreseeable future, unless it produces a radical shift in our understanding of ourselves, it raises no new ethical or legal challenges or dilemmas. That is, when the findings of the new science raise familiar ethical or legal problems, the resources to deal with them are already at hand. Moreover, nothing in the empirical literature yet *per se* compels changes in normative ethical or legal methods or conclusions. There are well informed, well-meaning people who believe that the latter claim is false and that what we know already compels normative changes. An implication of this entry, however, is that the burden of persuasion is on proponents of neuroscientifically-motivated changes to demonstrate that the findings that allegedly compel changes are sufficiently well-established and that the normative implications of such findings are sufficiently clear and uncontroversial to make such changes.

In neurolaw, the situation is especially complicated. What legislators, administrative officers, judges, and juries decide is authoritatively backed by the coercive power of the state, so the issues are not just intellectual and theoretical. They have more direct real world impact, even if the justification for this impact may be weak or even non-existent. For example, advocates, especially for the defense in criminal cases, have continuously and increasingly sought to admit neuroimaging evidence to bolster various defensive claims. Courts have sometimes been willing to admit such evidence, especially if the standard for admission is low, as it is in capital sentencing (*Lockett v. Ohio*, 1978). Thus, in a sense, the age of practical neurolaw may have begun, albeit largely prematurely. The findings of neuroscience must be translatable into the folk psychological categories of the law if they are to have relevance, but at present, neuroscience has little to offer to legal doctrine, policy and practice. In the future, one can expect that as the science matures and accumulates, it will make modest contributions to law, but it is unlikely to revolutionize law or to *per se* entail any particular changes in law.

II. The Sources of Inflated Claims for Neuroscience

Law and ethics have considered the findings from many sciences, including, sociology, different types of psychology, such as behaviorism and psychodynamic psychology, psychiatry, genetics, and now neuroscience. Although there are ethical and legal subdisciplines that have arisen as a result of the sciences, such as bioethics or psychiatric ethics, for the most part, none of these has been based on a revolutionary approach to law or ethics. They primarily use familiar legal and ethical concepts to address traditional issues the new sciences might produce. For example, genomic information about individuals might raise acute privacy or human enhancement issues, but these are traditional questions. The most revolutionary claim arising from these sciences is typically the hoary claim that determinism is incompatible with free will and responsibility. Each of the various sciences has presented itself as the newest proof of determinism that allegedly should upend doctrines and practices based on personal responsibility, typically in favor of one form or another of consequentially-based social control that is often mischaracterized as “medical” (Menninger, 1968). Nonetheless, none of these has engendered the type of academic and public enthusiasm (and fear) that neuroscience has produced. The question is why.

The relation of the brain to the mind and action has been at the center philosophical and scientific attention for centuries. We can roughly date the “neuroscientific” approach to understanding behavior to the case of Phineas Gage, a railroad construction foreman who suffered a severe injury to his

frontal cortex in 1848 as a result of an accident, but who miraculously lived. The traditional narrative, about which there is some doubt, is that prior to the accident Gage was a model of probity and rectitude, but after the injury he became disinhibited and his prior executive control skills deteriorated. Today we have a better understanding of the relation of frontal cortical function to executive control, but even then, the case was a powerful demonstration of the relation of brain structure and function to behavior. Not until the advent of non-invasive functional magnetic resonance (fMRI) imaging in the early 1990s, however, and not really until the early aughts, when scanners (often colloquially referred to as “magnets”) became more widely available, was a technology available that could investigate large numbers of non-clinical subjects. As a result of the increasing availability of fMRI, there is now an immense and increasing literature on the relation of brain to behavior that has fueled the scientific and popular imagination. This work seems somehow more rigorously scientific than previous sciences of behavior and the images produced, which are not “pictures” of the brain, can be ravishingly arresting. In a metaphor that seems question-begging because it assumes a form of mind/brain reductionism that is philosophically controversial, many enthusiasts claim we can now “look under the hood” of the acting agent to discern what the driving mechanisms are. Again, of course the brain is necessary for mind and action and we are discovering neural correlates and sometimes causes of mental states and actions, but acting human beings are usually not thought to be mere mechanisms like automobiles. Although, as the next section of the entry suggests, such beliefs are at present unjustified, the possibility has created great expectations.

I speculate that there are three sources of what I have termed neuroexuberance among philosophers, lawyers and others. The history of normative ethics and law as action guiding is overwhelming one of conflict and irresolution with no method to establish an obviously right answer (although many of course do believe that their position is the right answer). There is no experiment, even in principle, to indicate that humans *should* behave one way or another. It is all contestable. Nonetheless, many seem to believe that the findings of the “hard” science of neuroscience may hold a key. Even the Supreme Court of the United States fell prey to this belief when it distinguished neuroscience from social sciences (*Miller v. Alabama*, 2012, n. 5).

Second, many philosophers and lawyers are profoundly skeptical of deontology and especially of retributive justifications for state blame and punishment. Some incorrectly think that neuroscience proves that determinism is true, which, when coupled with hard determinist metaphysics

provides the desired conclusions that no one is really responsible for any behavior and that we should jettison outmoded and unjust retributively-based responsibility practices for consequentially-based social control. As noted, this argument has been made previously based on other behavioral sciences, but again, neuroscience seems like a more “real” science that at last will provide a genuine scientific basis for the argument. Last, behavioral neuroscience is inherently interesting and fun, albeit often hard, to perform. It provides a tangible result, not just an “argument” to which some other clever philosopher or lawyer will find a damaging and perhaps even decisive riposte. It thus offers an engaging and welcome respite from the common frustrations and annoyances of normative work.

Again, the above is speculation, but the amount of unjustified overclaiming and exuberance that contemporary neuroscience has produced is striking and cries out for an explanation. I have no stake in mine and would invite readers to speculate for themselves. I doubt that anyone will rigorously investigate the question.

III. The Present Limits of Neuroscience

Most generally, the relation of brain, mind, and action is one of the hardest problems in all science. We have no idea how the brain enables the mind or how action is possible (McHugh and Slavney, 1998, pp. 11–12; Adolphs, 2015, p. 175). The brain-mind-action relation is a mystery not because it is inherently not subject to scientific explanation, but because the problem is so hard. For example, we would like to know the difference between a neuromuscular spasm and intentionally moving one’s arm in exactly the same way. The former is a purely mechanical motion, whereas the latter is an action, but we cannot explain the difference between the two. Wittgenstein, famously asked: “Let us not forget this: when ‘I raise my arm’, my arm goes up. And the problem arises: what is left over if I subtract the fact that my arm goes up from the fact that I raise my arm? (Wittgenstein, 1953, ¶ 621.)” We know that a functioning brain is a necessary condition for having mental states and for acting. After all, if your brain is dead, you have no mental states and are not acting. Still, we do not know how mental states and action are caused. Wittgenstein’s question cannot be answered yet.

Despite the astonishing advances in neuroimaging and other neuroscientific methods—especially in understanding sensory systems such as vision and memory, for examples-- we still do not have sophisticated causal knowledge of how the brain works generally, and we have little information that is directly or even indirectly morally or legally relevant. The scientific problems are fearsomely difficult. Only in the present century have researchers begun to accumulate much data from fMRI imaging. New methodological problems are

constantly being discovered (E.g.,Bennett, C.M., Wolford, G.L., and Miller, M.B. 2009; Vul, Winkiekman, and Pashler, 2009; Button et al, 2013; Eklund, Nichols and Knutsson 2016). This is not surprising given how new the science is. Moreover, virtually no studies have been performed to address specifically normatively relevant questions. Ethics and law should not expect too much of a young science that uses new technologies to investigate some of the most intrinsically difficult problems in science and that does not directly address questions of normative interest. Caution is warranted, although many would think the argument of this entry is too cautious.

Further, neuroscience is insufficiently developed to detect specific, legally relevant mental content or to provide a sufficiently accurate diagnostic marker for even a severe mental disorder (Frances, 2009; Morse and Newsome, 2013, pp. 150, 159-160, 167). Many studies do find differences between patients with mental disorders and controls, but the differences are too small to be used diagnostically and publication bias may have inflated the number of such positive studies (Ioannidis, 2011). There are limited exceptions for some genetic disorders that are diagnosed using genomic information or some well-characterized neurological disorders such as epilepsy that is definitively diagnosed using electroencephalography (EEG), but these are not the types of techniques that are central to the new neuroscience based primarily on imaging. Indeed, when the American Psychiatric Association published its most recent version of the authoritative Diagnostic and Statistical Manual of Mental Disorders—Fifth Edition (DSM-5) in 2013, it conceded that no validated neurological diagnostic markers for major mental disorders such as schizophrenia and major affective disorder had been identified. Nothing has changed since then (Rego, 2016, but claiming that dementias may be an exception).

Nonetheless, certain aspects of neural structure and function that bear on legally relevant capacities, such as the capacity for rationality and control, may be temporally stable in general or in individual cases. If they are, neuroevidence may permit a reasonably valid retrospective inference about the defendant's rational and control capacities and their impact on criminal behavior. Some legal questions, such as whether a defendant is competent and what the agent will do in the future, depend on current rather than retrospective evaluation of the agent. Such evaluations will be easier than retrospective evaluation. Nonetheless, both types of evaluation will depend on the existence of adequate neuroscience to aid such evaluations. With the exception of a few well-characterized medical disorders, such as epilepsy, we currently lack such science (Morse and Newsome, 2013), but future research may provide the necessary data.

Let us consider the specific grounds for modesty about the current achievements of cognitive, affective, and social neuroscience, the sub-disciplines most relevant to ethics and law. fMRI is still a rather blunt instrument to measure brain functioning. It measures the amount of oxygenated blood that is flowing to a specific region of the brain (the BOLD—blood oxygen dependent level—signal), which is a proxy for the amount of activation that is occurring in that region above or below base-line activation (the brain is always and everywhere physiologically active). There is good reason to believe the BOLD signal is a good proxy, but it is only a proxy. The time lag between alleged activation and measurement and its spatial resolution are less than optimal (Roskies, 2013). These difficulties will surely be ameliorated by technological advances, but studies to date, especially if they used lower power scanners, do suffer from these limitations.

There are research design difficulties. It is extraordinarily difficult to control for all conceivable artifacts, that is, other variables that may also produce a similar result. Consequently, there are often problems of over-inference.

At present, most neuroscience studies on human beings involve small numbers of subjects, which makes it difficult to achieve statistically significant results and which undermines the validity of significant findings (Button et al., 2013; Szucs and Ioannidis, 2016). This phenomenon will change as the cost of scanning decreases and future studies will have more statistical power, but this is still a major problem. Most of the studies in cognitive, affective and social neuroscience have been done on college and university students, who are hardly a random sample of the population generally.

Many of the studies use other animals, such a rats or primates, as subjects. Although the complexity and operation of the neural structure and function of other animals may be on a continuum with those of human beings and there may be complete similarity at some level, there is reason to question the applicability of the neuroscience of behavior of other animals to humans. The human brain is capable of language and rationality that mark an immense difference between humans and other animals. To the best of our knowledge, other animals do not act for and are not responsive to reasons in the full-blown sense that intact human beings are. Is so-called altruistic behavior in orangutans, for example, the same as altruistic behavior in humans? Although the point should not be overstated, we should be cautious about extrapolating to human action from the neuroscience of the behavior of other animals.

Most studies average the neurodata over the subjects, and the average finding may not accurately describe the brain structure or function of any actual subject in the study. This leads to a more general problem about the applicability of scientific findings from group data to an individual subject, a

problem called G2i for “group to individual.” (Faigman, Monahan, and Slobogin, 2014) Scientists are interested in how the world works and produce general information. Law is often concerned with individual cases, and it is difficult to know how properly to apply relevant group data. For example, a neuroscience study that reports increased activation in some brain region of interest bases its conclusion on averaging the activation across all the subjects, but no subject’s brain may have activated precisely in the area identified. If such group data are permitted, as they now are for functions such as predictions, the question is how to use probabilistic data to answer what is often a binary question, such as whether to release a prisoner to parole because he is deemed no longer a danger to society. This is a topic under intensive investigation at present, and I assume progress will be made.

A serious question is whether findings based on subjects’ behavior and brain activity in a scanner would apply to real-world situations. This is known as the problem of “ecological validity.” Does a subject’s performance in a laboratory while being scanned on an executive function task that *inter alia* allegedly measures the ability to control impulses really predict the person’s ability to resist criminal offending, for example?

Replications are few, which is especially important for any discipline, such as law, that has public policy implications (Chin, 2014). Policy and adjudication should not be influenced by findings that are insufficiently established, and replications of findings are crucial to our confidence in a result, especially given the problem of publication bias (Ioannidis, 2011) and reproducibility skepticism (Chin, 2014; Open Science Collaboration, 2015). Indeed, replications are so few in this young science and the power is so low that one should be wary of the ultimate validity of many results. Indeed, a recent analysis by Szucs and Ioannidis (2016) suggests that more than 50 percent of cognitive neuroscience studies may be invalid and not reproducible. Drawing extended inferences from findings is especially unwarranted at present. If there are numerous studies of various types that seem valid, all converge on a similar finding, and there is theoretical reason to believe they should be consistent, then lack of replication of any one of them may not present such a large problem. The adolescent behavior example given in this entry’s introduction is a good example. But such examples are at present few, especially in legally and morally relevant neuroscience.

The neuroscience of cognition and interpersonal behavior is largely in its infancy and what is known is quite coarse-grained and correlational, rather than fine-grained and causal (Miller, 2010). What is being investigated is an association between a condition or a task in the scanner and brain activity. These studies do not demonstrate that the brain activity is a sensitive diagnostic

marker for the condition or either a necessary, sufficient, or predisposing causal condition for the behavioral task that is being done in the scanner. Any language that suggests otherwise—such as claiming that some brain region is the neural substrate for the behavior—is simply not justifiable based on the methodology of most studies. Such inferences are only justified if everything else in the brain remained constant, which is seldom the case (Adolphs, 2015). Moreover, activity in the same region may be associated with diametrically opposite behavioral phenomena—for example, love and hate.

Ethics and law are concerned with human mental states and actions. What is the relevance of neuroscientific evidence to decision making concerning human behavior? If the behavioral data are not clear, then the potential contribution of neuroscience is large. Unfortunately, it is in just such cases that neuroscience at present is not likely to be of much help. I term the reason for this the “clear cut” problem (Morse, 2011). Virtually all neuroscience studies of potential interest to the law involve some behavior that has already been identified as of interest, and the point of the study is to identify that behavior’s neural correlates. Neuroscientists do not go on general “fishing” expeditions (But see Bennett et al., 2009, for an amusing exception). There is usually some bit of behavior—such as addiction, schizophrenia or impulsivity—that investigators would like to understand better by investigating its neural correlates. To do this properly presupposes that the researchers have already well-characterized and validated the behavior under neuroscientific investigation. This is why, as the introduction claimed, cognitive, social and affective neuroscience is inevitably embedded in a matrix involving allied sciences such as cognitive science and psychology. Thus, neurodata can be no more valid than the behavior with which it is correlated. In such cases, the neural markers might be quite sensitive to the already clearly identified behaviors precisely because the behavior is so clear. Less clear behavior is simply not studied, or the overlap in data about less clear behavior is greater between experimental and comparison subjects. Thus, the neural markers of clear cases will provide little guidance to resolve behaviorally ambiguous cases of relevant behavior, and they are unnecessary if the behavior is sufficiently clear.

On occasion, the neuroscience might suggest that the behavior is not well-characterized or is neurally indistinguishable from other, seemingly different behavior. In general, however, the existence of relevant behavior will already be apparent before the neuroscientific investigation is begun. For example, some people are grossly out of touch with reality. If, as a result, they do not understand right from wrong, we excuse them because they lack such knowledge. We might learn a great deal about the neural correlates of such

psychological abnormalities. But we already knew without neuroscientific data that these abnormalities existed, and we had a firm view of their normative significance. In the future, however, we may learn more about the causal link between the brain and behavior, and studies may be devised that are more directly legally relevant. Indeed, my best hope is that neuroscience and the behavioral science will each richly inform the other and perhaps help reach what I term a conceptual-empirical equilibrium in some areas. I suspect that we are unlikely to make substantial progress with neural assessment of mental content, but we are likely to learn more about capacities that will bear on excuse or mitigation.

Here is an example of the current limitations of neuroscience for normative conclusions. A neuroscientist and I reviewed all the behavioral neuroscience that might possibly be relevant to criminal law adjudication and policy. With the exception of a few already well-characterized medical conditions, such as epilepsy, our review found virtually no solid neuroscience findings that were yet relevant (Morse and Newsome, 2013). Similar conclusions were reached after reviews of “brain reading” studies (e.g., “neural lie detection”) (Greely, 2013) and the addictions (Husak and Murphy, 2013). These conclusions are unsurprising. Behavioral neuroscience is a new discipline that is working on problems of immense conceptual and scientific complexity. Future conceptual and technological advances will certainly improve our knowledge base, but for now modesty is in order about what neuroscience can teach us about normative ethics or law.

Let us conclude this section with an observation that will always be germane even if neuroscience makes huge leaps forward. Neuroscience is a purely mechanistic science. Neurons, neural networks and the connectome do not have reasons. They have no aspirations, no sense of past, present and future. These are properties of agents. Ethics and law are addressed to agents. Thus, there will always be a problem of translation between the pure mechanism of neuroscience and the folk psychology of ethics and law. This is a greater problem for neuroscience than, say, for psychiatry and psychology. The latter sometimes treat people as mechanisms but also treat them as agents. Thus, they are in part folk psychological and the translation will be easier than for neuroscience. It is the task of those doing normative neuroethics and neurolaw always to explain precisely how neuroscientific findings, assuming that they are valid, are relevant to an ethical or legal issue. No hand waving is allowed.

IV. The Radical Challenge to Agency

This section addresses the claim and hope raised earlier that neuroscience will cause a paradigm shift in criminal responsibility and related doctrines and practices by demonstrating that we are “merely victims of neuronal

circumstances” or simply “packs of neurons” (or some similar claim that denies human agency). Fueled also by work in psychology (e.g., Wegner, 2002), this claim holds that we are not the kinds of intentional creatures we think we are. If our mental state, such as conscious decisions and intentions, play no role in our behavior and some or all are simply epiphenomenal, then traditional notions of responsibility, competence and the like that are based on mental states and on actions guided by mental states would be imperiled. But is the rich explanatory apparatus of intentionality simply a post hoc rationalization that the brains of hapless homo sapiens construct to explain what their brains have already done? Will our lives together be profoundly altered? Will ethical notions and the criminal justice system as we know it wither away as outmoded relics of a prescientific and cruel age? If we are just victims of neuronal circumstances, how should we live together?

Before continuing, we must understand that this is not the familiar challenge from determinism that can be answered by compatibilist metaphysics, which in one form or another holds that sufficient freedom of will and responsibility are possible even if determinism or something quite like it is true (a position discussed in detail in section V.A. below). Compatibilism does not save agency if the radical claim is true. If determinism is true, two states of the world concerning agency are possible: agency exists or it does not. Compatibilism assumes that agency is true because it holds that agents can be responsible in a determinist universe. It thus essentially begs the question against the radical claim. If the radical claim is true, then compatibilism is false because no responsibility is possible if we are not agents. It is an incoherent notion to have genuine responsibility without agency. The question is whether the radical claim is true.

Given how little we know about the brain-mind and brain-mind-action connections, to claim that we should radically change our conceptions of ourselves and our legal doctrines and practices based on neuroscience is a form of “neuroarrogance.” Although we may continue to see inflated claims and more numerous attempts to introduce neuroevidence in legal cases, the current state of neuroscience does not remotely prove that we are not agents (Morse, 2015; Moore, 2012; Mele, 2009, 2014).

The primary support in neuroscience for the radical claim was the work of neuroscientist, Benjamin Libet, and others who pursued similar work with similar and sometimes more striking findings (Libet, 1999; Soon et al., 2008). Let us focus on Libet’s work because he was the pioneer and it received the most attention. It is an excellent case study because no set of neuroscientific findings has generated as many claims about the normative implications of neuroscience. Indeed, it is perhaps the only body of research in neuroscience

that has received book-length treatment by philosophers and legal scholars concerned about its moral and legal implications (Mele, 2009; Sinnott-Armstrong and Nadel, 2010).

In Libet's work, subjects attached to an electroencephalogram (EEG) that measures electrical activity in the brain, were instructed to move a finger whenever they felt like doing so and to note by looking at a very precise clock when they first were aware of the urge/desire/impulse (there is dispute about how to characterize the subjects' mental state) to move their finger. Libet found that there was electrical activity in the supplementary motor area (SMA) of the brain, a "readiness potential," about 350-400 milliseconds prior to the subjects becoming aware of the urge to move and about 550 milliseconds before they actually moved. Libet and many others drew the conclusion that the brain activity fully causally explained the subjects' actions. Conscious decisions (i.e., the conscious intention to move) were apparently epiphenomenal. (Libet later tried to find what was waggishly termed "free won't" in the subjects' ability to "veto" the intention to move during the last 150 milliseconds, but this was a conceptual error on Libet's own account.)

Conceptual and empirical work seems to have exploded these claims (Mele, 2009; Mele 2014; Moore 2012; Nachev and Hacker, 2015; Schurger and Uithol, 2015). The function of the SMA is not well-understood, the existence of prior brain activity is unsurprising, and there is no reason to infer that mental states such as desires and intentions played no contributory causal role. It is hard to think of more trivial behavior so divorced from the agent's reasoning, whereas ethical and legal issues always involve reasons. It is not clear that the finding would hold for more complex, reason-guided behavior. Further empirical work has cast scientific doubt on the validity of Libet's explanation for the observed phenomena. And there is good evidence from psychology that mental states play a causal role in behavior. Perhaps most important, the radical claim violates ordinary experience and commonsense. Any proponent of such a case bears an enormous burden of persuasion that cannot possibly be satisfied at present. Nothing in neuroscience (or psychology) demonstrates empirically that we are not agents. It is possible that we are not agents, but as Jerry Fodor has argued, if we are wrong about the causal role of desire/belief/intent psychology, that is the wrongest we will be about anything since the belief in the supernatural (Fodor 1987, p. xii).

The radical view also entails no positive agenda. If the truth of pure mechanism is a premise in deciding what to do, no particular moral, legal, or political conclusions follow from it (Berman, 2008, first suggested this line of thought). The radical view provides no guide as to how one should live or how one should respond to the truth of reductive mechanism. Normativity depends

on reason, and thus the radical view is normatively inert. Reasons are mental states. If reasons do not matter, then we have no reason to adopt any particular morals, politics, or legal rules or to do anything at all.

Suppose we were convinced by the mechanistic view that we are not intentional, rational agents after all (and what would it mean for an agent to be “convinced” by data and argument if the radical claim is true?). If it is really “true” that we do not have mental states or, slightly more plausibly, that our mental states are epiphenomenal and play no role in the causation of our actions, what should we do now? If it is true, we know that it is an illusion to think that our deliberations and intentions have any causal efficacy in the world. We also know, however, that we experience sensations—such as pleasure and pain—and care about what happens to us and to the world. We cannot just sit quietly and wait for our brains to activate, for determinism to happen. We must, and will, deliberate and act.

Even if we still thought that the radical view was correct and standard notions of genuine moral responsibility and desert were therefore impossible, we might still believe that the law would not necessarily have to give up the concept of incentives. Indeed, Greene and Cohen concede that we would have to keep punishing people for practical purposes (although the term, “punishment,” which as moral valence, seems inappropriate in a world without responsibility). Such an account would be consistent with “black box” accounts of economic incentives that simply depend on the relation between inputs and outputs without considering the mind as a mediator between the two. For those who believe that a thoroughly naturalized account of human behavior entails complete consequentialism, this conclusion might be welcomed.

On the other hand, this view seems to entail the same internal contradiction just explored. What is the nature of the agent that is discovering the laws governing how incentives shape behavior? Could understanding and providing incentives via social norms and legal rules simply be epiphenomenal interpretations of what the brain has already done? How do we decide which behaviors to reinforce positively or negatively? What role does reason—a property of thoughts and agents, not a property of brains—play in this decision?

The radical claim is almost certainly false and provides no guidance for how we should live together. There is still a great deal of work to do for normative ethics and law.

V. Normative Neuroethics/Neurolaw

Neuroethics and neurolaw will be treated together because the same issues are important for both and ethics and law bleed into one another. Normative neuroethics and neurolaw are the application of traditional thinking to mostly

familiar problems that are also raised or are raised acutely by neuroscience. This section will therefore address a series of issues that have been most widely discussed: the “problem” of responsibility, enhancement of normal functioning, threats to liberty, competence, informed consent, end of life issues, and the ethics of caution.

A. The “Problem” of Responsibility

This section begins by addressing the most general alleged threat to responsibility: neurodeterminism. Then it addresses the criteria for responsibility before turning to the relevance of neuroscientific data to those criteria.

Does the new neuroscience in fact pose a threat to responsibility because it demonstrates that determinism is true and determinism is inconsistent with responsibility? This is entirely familiar ground for philosophers of responsibility. For over 2000 years, western thought has debated whether free will and responsibility are possible if determinism, universal causation or the like is true. The deterministic explanations have shifted with changes in theological and scientific understanding and fashion. God’s foreknowledge, social structure, unconscious psychodynamics, behavioral psychology, and genetics have all been seen as the basis for determinist understanding. Neuroscience is simply the newest alleged source of determinism on the block. Despite such changes, the alleged incompatibility between determinism and responsibility is an ancient issue. In this debate, free will is usually understood as the ability of people to act uncaused by anything other than themselves. If people do not have this ability, it is claimed, responsibility and other worthy goods such as autonomy may be unjustified. This thought is what disturbs people about scientific understanding of human behavior, which relentlessly exposes the numerous causal variables that seem to toss us about like light ships in a raging sea storm. Neuroscience, it seems, will finally support this challenge because it exposes that the brain, the final pathway to action, is nothing but a mechanism.

Neuroscientific or other biological causes such as genetic causation pose no more challenge to responsibility than non-biological or social causes. As a conceptual matter, we have no more control over social causal variables than over biological causal variables. In a world of universal causation or determinism, causal mechanisms are indistinguishable in this respect and biological causation creates no greater threat to our life hopes than social causation. For purposes of the free will debate, a cause is just a cause, whether it is biological, psychological, sociological, or as is usually the case with human behavior, some combination of all three.

There is no uncontroversial definition of determinism and we will never be able to confirm that it is true or not. As a working definition, however, let us assume, roughly, that all events have causes that operate according to the physical laws of the universe and that were themselves caused by those same laws operating on prior states of the universe in a continuous thread of causation going back to the first state. Modern physics teaches, of course, that there are indeterministically caused events in the universe, especially at the subatomic level. A few philosophers (e.g., Kane, 1998) utilize this as the basis for libertarian freedom of the will. But if indeterministic processes in the brain are part of the causation of behavior, this hardly seems to secure the type of freedom that we care about. After all, if the brain is in part a random-number generator, this does not seem to provide the agentic authorship of our actions that underpins our notions of responsibility. Thus, even if the original working definition of determinism is too strong, the universe seems sufficiently regular and lawful that it appears that we must adopt the hypothesis that universal causation is approximately correct. The philosopher, Galen Strawson, calls this the “realism constraint,” (Strawson, 1989) and it is certainly a view accepted by most scientifically-informed people, even if they are also humanists. If this is true, the people we are and the actions we perform have been caused by a chain of causation over which we had no control and for which we could not possibly be responsible. How would responsibility be possible for action or anything else in such a universe (see Cashmore, 2010, for a particularly strong argument)?

This is an “all or none” debate. If determinism or universal causation is true and incompatible with responsibility, then no one can be responsible for anything. Thus, unless there is a plausible answer either to the truth of determinism or to the alleged incompatibility of determinism with responsibility, genuine responsibility is impossible. At most we can have “as if,” simulacrum responsibility that is used to shape behavior but does not mean that people truly deserve praise and blame, reward and punishment, or autonomy or constraint. If this is the best solution possible, then there is still a real question whether it would “work” if everyone knew that holding each other responsible, praising and blaming, and rewarding and punishing had no adequate justification other than as incentives to shape behavior.

The notion that human beings have the god-like ability to act uncaused by anything other than themselves, is considered by most philosophers to be a “panicky” metaphysics (Strawson, 1982). Does this mean, however, that we must accept that responsibility is impossible? Within the philosophy of responsibility, there is a plausible, mainstream position termed “compatibilism” that holds that genuine responsibility is not inconsistent with

the truth of determinism and the absence of contra-causal freedom. For those who adopt some variant of this position, agents may be responsible if, roughly, they act intentionally, with reasonably integrated consciousness, suffer from no major rationality defects, and act free of compulsion. Compatibilists believe that this is sufficient “freedom of the will” to ground responsibility.

There are no decisive, analytically incontrovertible arguments to resolve the metaphysical question of the relation between determinism, free will and responsibility. And the question is metaphysical, not scientific. Nevertheless, compatibilism is an entirely plausible stance—indeed, in one variant or another it is the predominant view among philosophers of responsibility—and it is entirely consistent with moral and legal responsibility practices that now exist. After all, even if determinism is true, some people are rational and some people are not. Some people act under compulsion, such as in response to the threat of death, and (thank goodness), most people do not. Note again, however, if the radical claim that we are not agents is true, then compatibilism cannot save genuine responsibility because rationality and compulsion are normative notions that apply to agents.

In short, determinism and causation, whether arising from neuroscience or other causal explanations of human behavior, have nothing to do with actual moral or legal responsibility practices. Lack of causation or the falsity of determinism are neither criterial for nor foundational for responsibility (Morse, 2007). Using such terms simply confuses responsibility. Therefore, in principle, no amount of increased causal understanding of behavior, from neuroscience or any type of science, threatens the law’s notion of responsibility unless it shows definitively that we humans (or some subset of us) are not intentional, minimally rational creatures. And no information about biological or social causes shows this directly. It will have to be demonstrated behaviorally.

It is of course true that many people continue mistakenly to believe that causation, especially abnormal causation, is *per se* an excusing condition within our actual responsibility practices, but this is quite simply an analytic error that I have called “the fundamental psycholegal error (Morse, 1994).” It leads people to try to create a new excuse every time an allegedly valid new “syndrome” is discovered that is thought to play a role in behavior. Advocates cannot pick and choose their preferred causes without threatening all conceptions of responsibility. Causes *per se* excuse everyone or no one.

Now let us turn to the relevance of neuroscience to responsibility, beginning with a brief explanation of the meaning of this concept that is central to our morality and law.

Responsibility is a formal and informal ascription about agents that lead to specific types of judgments, such as whether the agent is blameworthy. The concept of responsibility in morality, law and ordinary interaction follows logically from the conception of the person and the nature of human interaction. Morality and law can only guide action if human beings are rational creatures who can understand the facts relevant to their situations and conform to rules and standards through intentional action. Responsible agents are therefore people who have the general capacity to grasp and be guided by good reason in particular contexts (Wallace, 1994). People, acting human agents, not brains and nervous systems, are and are not responsible. Responsibility, properly understood, has nothing to do with what most people understand by “free will.” Rationality, a behavioral criterion, is the primary touchstone of responsibility. Lack of compulsion is also a responsibility condition. Like rationality, it is also a behavioral criterion—essentially, acting in the absence of a very hard choice produced by a threat or acting in response to seemingly overwhelming desire. The latter is not well-understood, but it is part of ordinary parlance when we say that an agent cannot control himself.

Virtually all formal and informal responsibility criteria depend primarily on assessment of the agent’s rational capacities in the context in question. For example, a person is criminally responsible if the agent was capable of knowing the nature of his conduct or knowing that his conduct was wrong, a formulation first introduced in English law in *M’Naghten’s Case* (1843) and since adopted in one form or another in both common law and continental legal systems. Some people who commit crimes under the influence of mental disorder are excused from responsibility because their rationality was compromised, not because mental disorder played a causal role in explaining the conduct. The rationality criterion for responsibility is perfectly consistent with the facts—most adults are capable of minimal rationality virtually all the time—and with moral theories concerning fairness and justice that we have good reason to endorse.

The rationality requirement for responsibility--a general capacity for rationality in the context in question—is not uncontroversial and self-defining. It must be understood according to some normative notion of both what type and how much capacity is required. For example, legal responsibility might require the capability of understanding the reason for an applicable rule, as well as the rule's narrow behavior command. What rationality demands will of course differ across contexts. These are matters of moral, political and, ultimately, legal judgment, about which reasonable people can and do differ. These are normative issues, and, whatever the

outcome might be, the debate is about human action--intentional behavior guided by reasons.

Coercion or compulsion criteria for non-responsibility also exist, although they much less frequently provide an excusing condition. Properly understood, coercion obtains when the agent is placed through no fault of her own in a threatening “hard choice” situation from which she cannot readily escape and in which she yields to the threat. The classic example in criminal law is the excuse of duress, which requires that the agent must be threatened with death or serious bodily harm unless she commits the crime and that a person of “reasonable firmness” would have yielded to the threat. The agent has surely acted intentionally and rationally. The reason we excuse the coerced agent is not that determinism or causation is at work, for it always is. The genuine moral and legal justification is that requiring human beings not to yield to some threats is simply too much to ask of creatures like ourselves. Now, how hard the choice has to be is a moral, normative question that can vary across contexts. A compulsion excuse for crime might require a greater threat than a compulsion excuse for a contract. But in no case does compulsion have anything to do with the presence or absence of causation *per se*, contra-causal freedom, or “free will.”

A persistent, vexed question is how to assess the responsibility of people who seem to be acting in response to some inner compulsion, or, in more ordinary language, who seem to have trouble controlling themselves. Examples from psychopathology include impulse control disorders, addictions, and paraphilias (sexual disorders of desire). If people really have immense difficulty refraining from acting in certain ways through no fault of their own, this surely provides an appealing justification for mitigation or excuse. But what does it mean to say that an agent who is acting cannot control himself? People who act in response to such inner states as craving are intentional agents. A drug addict who seeks and uses to satisfy his craving does so intentionally. Simply because an abnormal biological variable played a causal role—and neuroscientific evidence frequently confirms this (e.g., Kalivas and Volkow, 2005)—does not *per se* mean the person could not control himself or had great difficulty doing so.

I believe that cases in which we want to say that a person cannot control himself and should be excused for that reason can be better explained on the basis of a rationality defect. In short, at certain times or under certain circumstances, the state of intense desire or the like make it supremely difficult for the agent to access reason. As always, causation and free will are not the issue. The assessment of human action in terms of rationality or common sense criteria such as “self-control” is in issue. Lack of control can

only be finally demonstrated behaviorally, by evaluating action. Although neuroscientific evidence may surely provide assistance in performing this evaluation, neuroscience could never tell us how much control ability is required for responsibility. That question is normative, moral, and, ultimately, legal.

Now let us turn to the relevance and implications of neuroscience for responsibility. The easiest case to address is when there is evidence of severely altered consciousness. Moral and legal responsibility require action (or intentional omission in cases in which the agent has a duty to act) and rationality, and in such instances either the person did not act because the definition of action requires reasonably intact consciousness or the action was not rational because rationality requires the potential for self-reflection that altered consciousness undermines. This does not mean of course that the responsible agent must be fully aware of all or even most of the causes of what he is doing. No one is, as a wealth of psychological studies has demonstrated. But the agent who may have no idea why he is behaving as he is will still be responsible as long as he is capable of being aware of what he is doing and potentially guidable by moral and legal rules and standards.

Neuroscience evidence might well be relevant to assessing the validity of the agent's claim that she did not act or our understanding of her mental state. Of course, the issue of the relevance of consciousness to responsibility was developed and people were able to evaluate such claims before any of the modern neuroscientific investigative techniques was invented. Neuroscience thus teaches us nothing new morally or legally about these cases, but it may well help us evaluate them more accurately. In most cases when neuroscience is relevant to the existence of action, there will typically be a well-characterized neurological abnormality, such as epilepsy, which involve traditional medical evidence rather than the new neuroscience. And, once again, neuroscience has not demonstrated that human beings generally are automatons rather than agents.

The more problematic cases are those in which the agent's consciousness was intact and he clearly acted, but there is nonetheless a question about the agent's responsibility, especially if there is evidence of some abnormality. For example, as sophisticated people understand, abnormalities do not cause violent conduct directly and they are not excusing conditions *per se* simply because they played a causal role. Instead, they produce behavioral states or traits, such as rage, impulsiveness or disinhibition, that predispose the agent to commit violent acts and that may be relevant to the agent's responsibility. After all, such states and traits can compromise rationality, making it more difficult for the agent to play by the

rules. For example, younger children and people with intellectual disability (formerly termed retardation or developmental disability) are not held fully responsible because it is recognized that their capacity for rationality is not fully developed. In these cases, once again, it is a rationality consideration, and not lack of free will, that is doing the work. Note, too, that if the capacity for rationality is compromised by non-biological causes, such as child-rearing practices, the same analysis holds. There is nothing special about biological causation.

Syndromes and other causes do not have excusing force unless they sufficiently diminish rationality in the context in question. In that case, it is diminished rationality that is the excusing condition, not the presence of any particular type of cause. For example, as the Supreme Court recognized (*Roper v. Simmons*, 543 U.S. 551 (2005)), adolescents who commit murder when they are 16 or 17 years old should not be subject to the death penalty because their capacity for rationality is not fully developed, and not because the process of myelination of cortical neurons is not complete. Incomplete myelination is only a part of the causal explanatory account of why the genuine mitigating condition—diminished rationality—existed (Morse, 2006).

Morality and the law were cognizant of the relevance of diminished rationality to responsibility and developed theories and doctrines of mitigation and excuse long before modern neuroscience emerged. But unless neuroscience demonstrates that no one is capable of minimal rationality—an implausible scenario—fundamental criteria for responsibility will be intact. On the other hand, neuroscience will surely discover much more about the types of conditions that can compromise rationality and thus may potentially lead to a broadening of current excusing and mitigating doctrines or to a widening of the class of people who can raise a colorable mitigating or excusing claim. Further, neuroscience may help adjudicate excusing and mitigating claims more accurately. At present, however, the findings of neuroscience are virtually never validly relevant to the actual evaluation of moral and legal responsibility (Morse and Newsome, 2013). The practical relevance will have to await future conceptual and scientific advances.

Most generally, some people think that executive capacity—the congeries of cognitive and emotional capacities that help to plan and regulate human behavior—is going to be the Holy Grail to help the law determine culpability. After all, there is an attractive moral and legal case that people with a substantial lack of these capacities are less culpable or competent. Perhaps neuroscience can provide specific data previously unavailable to identify executive capacity differences more precisely. There are two problems, however. First, significant

problems with executive capacity are readily apparent without testing, and criminal law, for example, simply will not adopt fine-grained culpability criteria. Second, the correlation between neuropsychological tests of executive capacity and actual real world behavior is not terribly strong (Barkley and Murphy, 2010). Only a small fraction of the variance is accounted for, and the scanning studies will use the types of tasks the psychological tests use. Consequently, we are far from able to use neuroscience accurately to assess non-obvious executive-capacity differences that are valid in real world contexts.

The last question concerning responsibility concerns possible disjuncts between behavioral and neuroscience evidence. The criteria for responsibility and competence are entirely behavioral, broadly understood to mean actions and accompanying mental states. But how should we respond if the agent is undoubtedly rational but the brain is abnormal, or there is clearly irrationality but the neurodiagnostic findings are unremarkable. In such cases, it is clear that actions speak louder than images (Mandavilli, 2006). Once again, it is people and not brains and nervous systems that are responsible or competent. At most, neurodiagnostic findings might be relevant to resolving unclear cases at the margin, but for various reasons at present, we lack the technology to accomplish this.

B. Enhancement of Normal Functions

The desirability and permissibility of permitting or even compelling access to enhancements of normal functions raises immensely difficult conceptual, moral, legal, political, and economic questions (Buchanan, Brock, Daniels, and Wikler, 2000, pp. 61-164, 181-203; Presidential Commission for the Study of Bioethical Issues, 2015 (including recommendations)). This section simply tries to touch on the major issues. What is interesting once again, however, is that although new scientific discoveries may raise the stakes, the questions raised about justice, equality, liberty, and efficiency are thoroughly familiar and rich theoretical resources already exist with which to address them.

Let us first make the controversial but plausible and necessarily simplifying assumption that we can identify a reasonable and workable conception of normality and abnormality that will apply relatively uncontroversially to a wide array of cases. Unless such a conception is possible, it will be impossible to distinguish between treatment and enhancement because that distinction is dependent upon a prior, baseline conception of normality/abnormality. The boundary between normality and abnormality can of course shift as conceptual understanding and empirical

data advance, but if the distinction is valid, then a treatment/enhancement distinction will also have force.

There is a lively debate in the literature about whether enhancements are wrong *per se*. Opponents such as Michael Sandel (2007) claim that enhancements threaten to undermine our essential humanity, whereas proponents such as Allen Buchanan (2011) believe that enhancements can contribute to human flourishing as long as they are properly regulated. Bioethicist Erik Parens (2005) believes the debate is somewhat overblown and formulates the debate in terms of authenticity, which both sides value but characterize differently. He argues that proponents of enhancement value self-creation whereas skeptics value self-discovery. Further, the proponents note that new technologies that can confer benefits can never be completely suppressed. Even if they are made illegal or considered wrong by large numbers of people, either a black market will be created if the enhancement is illegal or, if not, those who do not share the objection to enhancement will use them despite the disapproval of their neighbors.

We already permits a wide array of enhancements for those who can afford them. Some are quite expensive—such as purely cosmetic surgery in the absence of disfigurement, psychotropic drugs prescribed to make people without diagnosable disorder feel even better, and prep courses for standardized tests—and, consequently, their availability is limited to those with the resources to purchase them. Others, such as the use of caffeine or nicotine to enhance mental acuity, are quite inexpensive and thus available essentially to everyone. There is certainly no general presumption that enhancement is *per se* undesirable or immoral. The law regulates the sale and use of such enhancements very little or indirectly by requiring warning labels, prescriptions and other various means that scarcely prevent access for those with the necessary resources. Private preference, conscience and pocketbooks are thus the primary predictors of which people obtain which enhancements.

Some potentially enhancing agents are largely or entirely prohibited either generally, because the government has decided they are too dangerous for almost anyone to use them, such as certain stimulants, or in particular contexts, such as sporting events, in which use would be considered unfair or otherwise undesirable. Such limitations do not undermine the observation that enhancement by cognitive and biological techniques is already widely permissible and acceptable in our moral, political and legal culture. This outcome is not surprising in a society that values personal liberty and uses primarily market mechanisms to develop and distribute most goods.

The use of enhancements raises thorny questions of distributive justice when the enhancements substantially increase the possibility that the agent

will thereby obtain other, socially desirable goods, such as access to better schools, jobs or the like. Is it really fair, for example, that a student from a wealthy home who already has enormous educational advantages by going to better schools should then have the additional advantage of taking a prep course for the SAT or of having access to a prescription for substances that may increase alertness, concentration and other qualities that promote excellent performance on cognitive tasks? Many views of justice deny that this is fair because they hold that most inequality is not justified, but others endorse the inequalities that result as justified by liberty, efficiency and other values. We cannot resolve these issues, but we should note that as enhancements become more effective, the potential for unjust distribution will increase, especially if the original distribution of endowments and access to the enhancement are unfairly unequal.

The discoveries of neuroscience may well provide highly effective, precise enhancement possibilities that will affect physical and cognitive functions that strongly predispose to improved performance in important life tasks. Let us also assume that such enhancements would not have undesirable side effects. If so, and they are not freely available because they are too expensive for many citizens, then potentially unfair increases in inequality will result. This could be addressed by prohibition or by making the enhancement more freely available by subsidization or other mechanisms. The latter would not have the desired effect, however, because people do not have equal endowments to be enhanced. Unless, miraculously, an enhancement caused everyone to produce precisely the same performance, the whole performance distribution would simply shift upwards, but the original inequalities would remain although they might be reduced if, as seems to be the case, those with lesser endowments achieved greater gains than those with more. We should also note that highly effective enhancements may be used in the service of vice and not just for virtuous pursuits.

Enhancing everyone in certain ways may perhaps be socially desirable and should be implemented (by various inducement mechanisms) even if it does not reduce inequalities. For example, if people with low normal intelligence could enhance their cognitive abilities, then they and all of society might be better off, but such people would not become the cognitive equals of those better endowed *ex ante* if the latter were also permitted to use the same enhancement. It is also interesting to contemplate whether, to pursue greater equality, certain enhancements would be permissible only for those people whose normal abilities were below some threshold and prohibited for those above that threshold. I assume that in the United States such a scheme would be held unconstitutional at present as a denial of both liberty and equal

protection, but if certain inequalities threatened the social fabric, one can imagine a court upholding such a law. In more highly regulated legal orders, such laws may be less problematic.

May the State make enhancements obligatory? Some enhancements already are imposed. Public education or some equivalent is a requirement for all citizens because the state interest in promoting a citizenry capable of informed participation in the political process and economic productivity is extremely weighty. No liberty is absolutely protected and any may be infringed if the government purpose for infringement is sufficiently strong. A balance must always be struck. Suppose, for example, there was widespread agreement that general improvement in cognitive skills would be desirable for reasons similar to those justifying compulsory education. Why shouldn't everyone be compelled to accept a new, neuroscientifically-discovered, non-harmful enhancement for the good of the whole society? The liberty of those who would not wish to be enhanced would be infringed, but perhaps the infringement would be justified.

Consider the following analogy. Forced inoculation—and note that preventive inoculations are another form of enhancement—might be imposed on all citizens to avoid a dreadful infectious disease epidemic, including on those people who objected strongly on religious, moral or other grounds. The examples are distinguishable, of course. One might say that failure of some people to enhance themselves cognitively does not threaten to make society worse off; it just fails to make some people better off. In contrast, failure to inoculate threatens only to make society worse off. The distinction is genuine, but the baselines against which welfare are assessed are normative and shift easily. It would not be difficult to re-conceptualize refusal in the cognitive case as threatening harm. For example, a more communitarian society that expected citizens to exert their best efforts and to accept enhancements in order to increase the welfare of all would treat a person who refused to accept enhancement as a threat to the society. In sum, as the social benefits of an enhancement increase, the State interest in imposing it will also increase, but traditional concerns for liberty and freedom of thought and expression should politically and legally constrain compelled enhancement.

The widespread availability of effective enhancements could profoundly affect our conception of normality, raising the threshold of normality considerably. If this occurred, then certain abilities that were previously considered normal would now be considered abnormal and thus would qualify for treatment, not enhancement. If this occurs and the disadvantage of those below the normality threshold is substantial, then such people would have a strong justice claim for the State to provide such treatment if they cannot

afford it. According to virtually all current moral and political theory, the duty to provide treatment to the least well off is far greater than the duty, if any there be, to provide enhancement. And all these considerations are not just ethical musings. The Presidential Commission for the Study of Bioethics (2015) has already considered this issues and has made recommendations, including that enhancers not contribute to existing inequalities.

What is the potential threat of enhancement to our identity and humanity, to our very nature (e.g., Harris, 2007; Garreau, 2005; Sandel, 2007). When we use agents that affect energy, cognition and mood, we are different from our “base-rate,” but we usually remain our recognizable selves and none of us remotely approaches “perfection,” whatever that might mean. This is generally unproblematic for the reasons already discussed. But massive changes produced by hitherto unimaginable discoveries in neuroscience could écreate such discontinuities between our usual and enhanced selves that our identities and sense of what it means to be human might be compromised. The fear is that we would become mechanistic robots rather than real people, beings that are engineered rather than largely self-created. Many philosophers and others believe that solving the problem of consciousness is beyond the cognitive capacities of human beings. Suppose, however, that immense cognitive enhancements permitted us to solve the problem. Such a discovery would revolutionize our understanding of biology, would permit the invention of enormously powerful behavior control techniques, and would almost certainly profoundly alter our sense of ourselves and our moral and political beliefs and arrangements. The prospect of this brave new world terrifies many people who would like to put substantial limits on enhancements of this magnitude. The history of technology indicates, however, that technological advances that have clear moral implications—consider controversies concerning stem cell research, the use of steroids among professional athletes, or the use of massively destructive weapons—will be used if doing so seems morally justified or seems to confer a significant advantage. One can only imagine what the ethical debates of the future will be, but one can safely predict that powerful and safe behavior control techniques will not be successfully suppressed for long.

Finally, let us briefly consider two technologies that are currently used for therapy, but that might in the future provide knowledge that would lead to enhancements or to enhancements themselves: brain-computer interfaces (BCI) and deep-brain stimulation (DBS). BCI is a collaboration between a brain and a device that enables signals from the brain to direct some external activity, such as control of a cursor or a prosthetic limb. The interface enables a direct communications pathway between the brain and the object

to be controlled. This relatively recent and still experimental treatment has shown promise for helping patients such as those with neuromuscular disorders or stroke victims to communicate or to manipulate their environment (He et al, 2013), or for other purposes, such as determining whether a patient with a disorder of consciousness is conscious (Lulé et al, 2012). This technique is non-invasive and has no side-effects. DBS is a treatment method in which a small electrode is inserted in the brain that is attached to an externally-worn device that controls the dosage of the pulses of electricity to the targeted area. It is now an accepted treatment for Parkinson's Disease and it is entirely experimental for use with mental disorders such as refractory major affective disorder and obsessive-compulsive disorder (Holtzheimer and Mayberg, 2011). DBS is of course invasive and does have potentially serious side effects. The number of mental patients treated with DBS is very small and it is impossible to draw conclusions yet about its comparative efficacy and probability of risk of serious side effects.

The bioethical literature addressing such techniques is relatively sparse (see Schneider, Fins and Wolpaw, 2012 and Morse, 2012 for discussion and sources), but the primary concerns are informed consent (discussed generally in part V.E. below) and efficacy. I raise them here, however, because both hold the promise of enhanced understanding of the brain-behavior relation that might someday lead to enhancements.

In conclusion, it is worth noting yet again that neuroscience raises few new issues. Enhancement based on behavioral psychology and genetic manipulation are familiar and much discussed topics. Neuroscience simply raises the stakes by potentially providing more effective, targeted techniques for enhancement.

C. Potential Threats to Civil Liberties: Privacy, Prediction and Treatment

Neuroscientific discoveries may raise the specter of profound challenges to civil liberties that I will discuss under the rubrics of privacy, prediction and treatment. Other sciences, too, might make discoveries that would raise similar challenges so the following discussion surely generalizes. The potential of neuroscience to invade our privacy by revealing various aspects of our private, subjective experience and to undermine our autonomy by predicting and controlling our behavior without our consent may produce the strongest reaction against its use and substantial regulation. On the other hand, the use of techniques that permit genuinely accurate lie detection, control of dangerous conduct and other valuable ends may be so alluring that the temptation to use it will be great. One need only think about legal

responses to the “war on terror” to recognize that justifying the use of privacy-invasive techniques may not be so difficult after all.

What constitutional or legislative limits may be placed on such techniques? This will of course depend on the political and legal regime in which such techniques are considered for regulation. In the United States, for example, in a case that provides a technological analogy, the US Supreme Court held that the 4th Amendment prohibition against unreasonable searches and seizures barred police use of heat sensors from outside a private home to detect marijuana plants within (*Kyllo v. U.S.*, 533 U.S. 27 (2001)). If and when we are able to use brain states to infer mental content, what will happen to the privilege against self-incrimination that is so important in Anglo-American law? In the United States, testimonial evidence is privileged but so-called physical evidence is not. For example, the state may involuntarily use a breathalyzer to determine if a driver is drunk. As legal scholar Nita Farahany has argued, this distinction is confused, and she offers a more nuanced set of criteria, but the traditional distinction is still in use. Will brain states used to infer content be testimonial or physical? The issue is completely open and very important.

It appears that in most liberal societies (broadly conceived) the state will not be able to use neuroscientific investigative techniques to go on “mental fishing expeditions” generally, but various state interests may permit infringing hitherto protected interests. Neuroscience undoubtedly poses a threat to privacy because its techniques might be used to gather information about mental content. Again, this is not a new issue, but neuroscience may create increased concern because its techniques may be more invasive than previous technologies used for similar purposes.

Neuroscientific techniques might also increase the ability to make accurate predictions about various forms of future behavior (see, e.g., Aharoni et al., 2013; Pardini, Raine, Erickson, and Loeber, 2014). If some behaviors that are particularly socially problematic can be accurately predicted, especially in an era of “big data,” once again there will be a temptation to use such techniques for screening and intervention. For example, criminal and antisocial conduct are an immense social problem in the US; over two million people are incarcerated in state and federal prisons. Given the association between neuropsychological and neuropsychiatric abnormalities and criminal conduct and the increasing ability to detect such abnormalities, it is plausible to assume that neuroscientific techniques may well enhance the ability to predict future antisocial conduct among both those who have not yet engaged in such conduct and those who have. The social and personal costs of criminal conduct are so great that if the predictive techniques were sufficiently

sensitive and remedial intervention of any sort were possible, there would be a strong temptation to screen and intervene.

It would be far easier to justify screening and involuntary intervention among people otherwise justifiably under state control, such as involuntarily committed psychiatric patients and prisoners and others under criminal justice control. Although involuntary patients and prisoners have rights in all liberal societies, they may be curtailed and techniques that increased the accuracy of predictions of recidivism would probably be acceptable to promote public safety.

Widespread screening of apparently at-risk children and adolescents, or the general population—even if the risk status was identified by objective, valid measures--would be legally and politically fraught, especially if the predictive techniques and the necessary interventions were particularly invasive of liberty. Labeling and stigma effects and the potential for racial and ethnic bias would be frightening. The widespread usage of psychotropic medications such as methylphenidate among public school children suggests, however, that a screening/intervention scenario would not be unthinkable if predictive accuracy and remedial intervention were sufficiently successful and the “side effects” of both could be strictly limited. At present, the science is not sufficiently advanced, political resistance would be intense in most western societies, and, at least in the United States, it is probable that such schemes, even if adopted, would not survive a constitutional challenge. But it is difficult to envision how liberal societies would respond to techniques that accurately identified risk-creating variables and effectively intervened to prevent serious social and personal harms. Traditional notions of privacy and liberty might be changed considerably.

The new neuroscience increasingly is producing direct, biological intervention in the working of the brain and nervous system using techniques such as deep brain stimulation, vagus nerve stimulation and transcranial magnetic stimulation (Marangell, Martinez, Jurdi, & Zboyan, 2007). The potential for such methods not only to treat recognized disorders or to lead to enhancements, but also to change thoughts, feelings and actions, which is often polemically characterized as the potential for “mind control,” is particularly disquieting. Many consider this a greater threat to liberty than genetic intervention.

The government already has the authority to compel the use of psychotropic medications under relatively limited circumstances, and the failure of a patient to take needed medication that leads to dangerous conduct may be a source of criminal or civil liability even if the patient is not responsible when unmedicated. Nonetheless, the potential for widespread

intervention to change behavior is apparent. As is well known, the biological and behavioral definitions of abnormality and disorder can be controversial. At the extremes, of course, there is little problem, but the criteria for abnormal brain structure or function are not obviously self-defining, the criteria for behavioral abnormality are even more fluid, and there is a tendency to pathologize problematic behaviors and the structures and functions that seem associated. Thus, there is no guarantee that a relatively reasonable and workable criterion of abnormality will impose strict limits on the ability of the state to compel behavior-altering interventions.

For example, the US Supreme Court has decided that the State may involuntarily medicate a prisoner with psychotropic medication only if it is medically appropriate and necessary for the safety of the inmate or others in the prison. If these criteria are met, the prisoner's liberty interest in avoiding unwanted psychotropic medication must yield (*Washington v. Harper*, 494 U.S. 211 (1990)). Although the provision of safe conditions in prison is an important state interest, there is widespread agreement that medication cannot be used solely to control prisoners' behavior. Consequently, the concept of "medical appropriateness" is doing the justificatory work. In the case of a manifestly psychotic and consequently dangerous inmate (but most people with psychotic mentation are not dangerous) who refuses to consent to treatment, there would be little disagreement about the appropriateness of involuntary medication. Now, however, mounting evidence suggests that a class of antidepressant drugs with a relatively benign side-effect profile, the selective serotonin re-uptake inhibitors, may reduce the incidence of violence among prisoners who do not obviously meet the diagnostic criteria for a depressive disorder (*Walsh and Dinan*, 2001). It is extremely tempting to assume that many potentially violent prisoners have "underlying" or hidden depressive disorders or that the risk of violence is a pathology that is medically appropriate to treat. There is no incontrovertible conceptual or empirical block to making such assumptions. Therefore, it is possible that courts might approve a program that compelled medication after appropriate screening in order to serve the goals of safety and "treatment (*State v. Randall*, 532 N.W.2d 94, 106-110 (Wis. 1995))."

For another example, the US Supreme Court has held that under limited conditions the State had the right to medicate with psychotropic medication a psychotic criminal defendant solely for the purpose of restoring the defendant's competence to stand trial (*U.S. v. Sell*, 539 U.S. 166 (2003)). The State's interest in adjudicating guilt and innocence in cases of serious crimes of violence and property was deemed sufficient to warrant infringing the defendant's admitted liberty interest in deciding whether to take such

medication. The Court permitted such treatment only under limited conditions, but there will be inevitable pressure to use such medication or other techniques that may lead to a final determination of guilt or innocence in criminal cases.

A highly contentious, related issue, is the ethics of coercively providing or offering interventions to prisoners that might lead to early release on consequential grounds. Such cases could arise if a prisoner is not dangerous in prison, but suffers from some treatable condition that makes him a danger to the community. Consider pedophilic offenders, for example. There would be enormous civil liberties issues if the treatment were coercively imposed despite a competent prisoner's objection, but suppose the state offered early release in exchange for accepting the treatment. In the latter case, some theories of coercion suggest that this scenario is not coercive because the prisoner has no base-line normative right to be released early. Thus, the treatment proposal is really an offer, not a threat, and offers are held to increase freedom. Even if such offers are acceptable, early release might still offend retributive conceptions of justice that hold that responsible offenders should get their just deserts for past crimes.

The examples just given can of course be generalized. Once again, the state will have more power involuntarily to intervene in the lives of those it already controls, such as prisoners and patients, than in the lives of other citizens, but wider programs may be envisioned. Public health officials already pathologize violence, especially involving the use of guns, as a public health problem, and it is easy to imagine compelled treatment of the risk of violence as a justified method of protecting public health. Present involuntary outpatient commitment is usually limited to people with serious mental disorders, but as the example above indicates, adroit redefinitions of pathology and medical appropriateness might widen the state's net considerably. Again, the current science and political will to accomplish effective widespread behavior control are lacking. Nonetheless, as screening and intervention methods become more precise and effective, there will be pressure to use them and proponents will defend their legitimacy.

If neuroscience or other sciences ever reach the levels of understanding and efficacy necessary to make the foregoing civil liberties concerns a realistic possibility, it is difficult to predict what legislatures and courts will do. If there are pressing social problems that seem soluble by a technological fix, political and legal constraints may weaken, even in those societies that emphasize liberty more strongly.

D. Competence

Civil and criminal law have many doctrines concerning competence that may affect an agent's liberty, autonomy and other important interests. If a person is incompetent in a particular context, the usual rules governing that conduct are not applied. For example, in some cases a contract may be avoided by a party who was incompetent to contract or a will might not be given effect because the testator (the person who executed the will) was not competent to make a will. In criminal law, a defendant cannot be tried if he is incompetent or unfit to stand trial, cannot be sentenced if he is incompetent to be sentenced, and cannot be executed if he is incompetent at the time of execution. Roughly speaking, all competence doctrines are functional and depend on the agent's rational capacities in the context in question. A testator will be deemed incompetent if at the time of making the will he did not understand the nature and extent of his property and who his heirs were. A criminal defendant is incompetent to stand trial if he does not understand the nature of the proceedings and is not able rationally to assist counsel. How much rational capacity the agent must have to be deemed competent is a normative question that can vary across contexts. Neuroscience cannot tell us what the standards should be, but it may help thinking about such standards if neuroscience were to make substantial inroads in understanding the limits of humans' rational capacities.

At present, competence must be assessed behaviorally, focusing on the agents mental states in the context in question. These issues long antedated the advent of the neuroscience and the new neuroscience raises no new issues. The question neuroscience raises is whether it can help make these evaluations without infringing on other interests. At present the answer is no because the neuroscience is insufficiently advanced. Behaviorally clear cases will not require neuroscientific assistance and unclear cases will get little help as a result of the "clear cut" issue (discussed in section III. above). In the future, neuroscience may be able to help more because there may be clearer neural markers to help resolve close cases, but for now the issues must be resolved behaviorally.

E. Informed Consent

It is a commonplace that the legal doctrine of informed consent protects a patient's or research subject's liberty and autonomy. Competent adults have a right in virtually all circumstances to control what is done to or with their bodies and minds. There are controversies about how much information should be disclosed and what level of understanding a patient or research subject must achieve in order to make consent valid (Berg, Appelbaum, Lidz, and Parker, 2001), but all informed consent standards are based on the assumptions that the potential patient or subject is rational and that the

information provided will aid the person's ability to make a rational decision about her self-determination. Once again, morality and the law's model of the person as an intentional, rational agent grounds this doctrine. Exceptions to the need to obtain informed consent involve situations in which the person's autonomy interests are subjugated to other values, such as cases of compulsory treatment, or in which the person is not rational and cannot properly be restored to rationality, say, by medication or psychological treatment,, in which case a substitute decision maker will be required. A further complication is what values a substitute decision maker will use. Should the decision maker try to ascertain what this subject would so when rational (if the person has ever been sufficiently rational) or should an objective standard—what would a reasonable person do under the circumstances—be applied? There are arguments for both approaches.

Contemporary neuroscience raises at least two potential issues for the theory and practice of informed consent. The first is whether neuroscience can teach us anything new about the ability of people to process and to use information under various conditions, such as stress. Once again, the ultimate issue is behavioral—it is about a person's cognition rather than about the brain per se—but neuroscience will surely improve our understanding of information processing. Better understanding would be unlikely to alter the doctrine of informed consent profoundly unless it radically altered our model of the person. Indeed, most of the controversy about the requirements for informed consent and most legal developments have been produced by changing views about the moral issues, such as how much autonomy must be protected and balanced against other values, and not by scientific data about the brain or behavior. Nevertheless, better understanding of cognition might alter practice considerably.

The second issue concerns consent to neuroscientific research. Doctrines of informed consent to research developed somewhat independently of and parallel with informed consent to treatment, but the justification is the same and is perhaps more important because being a research subject often brings no potential benefit to the subject other than altruistic satisfaction or some compensation and it may impose substantial costs. Once again, improved understanding of brain function will not alter the fundamental doctrine and practice unless the model of the person changes, but neuroscience research does raise a number of important, interesting traditional issues.

Much research will be done on neurologically and psychiatrically impaired people, which raises difficult informed consent issues if the impairments affect the potential subject's rationality. Neuroscience may help

to identify those incapable of giving adequate informed consent to neuroscientific and other forms of research. People with abnormalities are prone to the “therapeutic misconception,” the error of believing that the research will benefit them, even when they are explicitly told that it will not or that it cannot be predicted that participation will help. This is a general problem about obtaining informed consent in a wide array of biomedical contexts.

Another traditional issue neuroscientific research raises is incidental findings, discoveries about the subject that were not being investigated (Wardlaw et al., 2015). For example, a structural brain scan to measure the volume of a region of interest may disclose some unexpected brain abnormality. There are estimates that as many as twenty percent of research scans reveal such abnormalities and about three percent warrant significant medical follow-up. Should subjects be informed in advance that such findings may be discovered? What should be the protocol for deciding which findings are significant? Should subjects be informed about apparently insignificant findings? If what the scan discovers implicates the interests of others because, for example, an abnormality might affect the subject’s job performance and risk harming others, should the experimenter have a duty not only to alert the subject but perhaps to alert the authorities in an appropriate case? Again, these are familiar issues and in the current climate favoring fuller subject autonomy and disclosure, more information should be provided *ex ante* and more disclosed after the scans have been read and interpreted.

The complexity of the brain and its relation to behavior and to one’s conception of the self raises somewhat speculative but profound issues. Biomedical research can potentially disclose threatening information, such as the presence of hitherto unrecognized disease or the potential for it. Neuroscience research in particular can arguably discover information about the brain that could alter one’s sense of self or that is especially invasive of the subject’s privacy. Further, if neuroscientific investigation becomes more invasive, the potential for unpredictable effects on behavior and personality would increase substantially. Indeed, the brain is so complicated that often we may be unable accurately to identify for the research subject the potential risks of neuroscience research. The ethical question then is familiar—how much uncertainty is tolerable—but the stakes may be much higher.

The informed consent to research issues just raised are traditional. I foresee no major changes in practice for neuroscientific research, but the application of existing rules and practices may be contextually altered. For example, if some of the speculative problems arose, I assume that either especially rigorous informed consent would be required or perhaps there

would be state regulation. At present, however, moral and political theory and the law have the resources to deal as well with neuroscience research as with any other type of human subjects research with similar benefit/cost profiles.

F. End of Life Issues

In the developed world, death is increasingly defined as brain death rather than the cessation of heart and lung function. As the result of disease or injury, many people have massive disorders of consciousness that are irreversible but are not the equivalent of death. People in what is termed the “persistent vegetative state” (PVS) are apparently unresponsive to stimuli and completely lack awareness, although they do have sleep/wake cycles. They can also be kept alive for very long periods with artificial nutrition and hydration.

A central question in bioethics is under what conditions it is justifiable to discontinue artificial life supports and simply to let the patient die. Again, this is a thoroughly familiar issue. The contribution of neuroscience may be to help decide if a patient is really in a PVS or is in what is called the minimally conscious state (MCS), in which there is awareness. There are now a number of studies using neuroimaging that suggest that some people diagnosed as in a PVS may be misdiagnosed and may in fact have awareness (Fins et al., 2008). If so, this complicates enormously the decision whether to discontinue life support. The patient in the PVS has little existence other than purely physiological, whereas the patient in the MCS may have a modicum of a psychological life and has some chance for recovery. PVS justifies discontinuing life supports far more readily than MCS (if anything does, which certain religious belief systems deny). These are difficult issues. Although neuroscience cannot tell us when discontinuing life support is justified and it cannot yet make definitive diagnoses of whether a patient is in the PVS or the MCS, in the future better diagnoses, whether based on neuroimaging or other techniques, will help make these decisions more rational because the facts involved will be clearer.

VI. Neuroevidence in the Criminal Law Courtroom

Quite recently we finally have preliminary data about how neuroscientific information is being used in criminal cases. Four very interesting empirical studies from the United States (Farahany, 2015), England & Wales (Catley and Claydon, 2015), Canada (Chandler, 2015), and the Netherlands (de Kogel and Westgeest, 2015) have attempted to discover the extent to which and in what way neuroscientific evidence is used in criminal cases. Recent excitement about the potential legal implications of non-invasive brain imaging by fMRI motivates this work. These studies begin

to examine the reality of neuroscientific influence in criminal cases. All focus on appellate cases reported in various data bases for somewhat different periods in the range of years from 2000-2012 and all are admirably cautious about the methodological limitations of the study sample. None purports to be an accurate representation of the use of neuroscientific evidence throughout the criminal justice system and other methodological quibbles may be raised, such as the failure to use independent inter-rater reliability for characterizing the cases. All use a very expansive definition of neuroscience that includes techniques and data that long antedate the new neuroscience. At most, the data are suggestive. Nonetheless, the studies are interesting and innovative.

The late, great baseball scientist, Yogi Berra, was apocryphally quoted as saying, “It’s déjà vu all over again.” The data indicate that the courts amake the classic mistakes about the relevance of neuroscience and behavioral genetics to criminal cases that have bedeviled the reception of behavioral science in general and of psychiatry and psychology in particular. The overarching classic mistake is misunderstanding or uncritically accepting the validity of apparently relevant science and misunderstanding the relevance of the science to the specific criminal law criterion at issue, which are primarily acts and mental states. There are no brain or nervous system criteria in criminal law for any doctrine. In particular, courts too often do not understand the following. Metaphysical free will is not a criterion for any criminal law doctrine and it is not even foundational for criminal responsibility in general. Causation in general and brain causation in particular, even causation by abnormal variables, are not *per se* a mitigating or excusing condition and causation *per se* is not the equivalent of compulsion, which is an excusing condition. And, finally, people with the same diagnosis or condition are behaviorally heterogeneous and ultimately it is the behavior that is legally relevant, not the diagnosis. In one form or another, most of these cases exhibit these mistakes and confusions. It is no surprise that one of the authors, Professor Nita Farahany, characterizes the cases as follows: “That use [of neurobiological research in criminal law] continues to be haphazard, ad hoc, and often ill conceived” (Farahany, 2015, pp. 488-89).

Not surprisingly, sentencing decisions were the most common context for the introduction of neuroscience evidence, but it was also used to resolve questions about many criminal responsibility doctrines and, surprisingly, competence, which as we have seen, is a functional behavioral determination. Perhaps the most striking finding is how infrequently the new neuroscience of functional imaging and related techniques are used. This varies across jurisdictions, but the large majority of cases involve the “old” neurology or

the old neuropsychology that uses classical structural imaging or behavioral methods to assess brain functioning associated with well-characterized neurological conditions, such as epilepsy and frontal lobe injuries or lesions. Such diagnostic methods are far more common than fMRI, and in the Dutch and Canadian samples, there is virtually no functional imaging evidence.

In sum, these studies suggest that the influence of the new neuro-investigative techniques applied to individual cases for forensic assessment is quite modest. Even when inferences are drawn in individual cases using group data about the consequences of various neurological conditions, the studies used are often classic behavioral studies rather than neuroimaging investigations. Indeed, careful examination of the expanded case studies the papers present indicate that in most instances the neuroscientific evidence was far less important than the behavioral evidence and the former was used largely to buttress the latter. The neuroevidence was rarely dispositive and in the other cases it is impossible to know from these papers' summaries of the case reports how influential the additive neuroevidence was.

The first question when considering the admissibility of scientific evidence, as always, is the degree to which the basis of the testimony has been established. We have already seen in Part III. of this entry that legally-relevant neuroscience is hardly well-established at present. As noted above, it is no critique of contemporary neuroscience to note that it is working on one of the hardest problems in science, the relation of the brain to mind and action. For a specific example, the apparently wide but not universal Dutch acceptance of a brain disease model of addiction that guides legal decision-making fails to confront the hard questions about the status of the science. Judges are not yet in a good position to evaluate neuroscience and may be either too critical or too uncritical (see Rakoff, 2016, for an analysis by a neuroscientifically-informed federal judge). In what follows, however, I shall assume that the science is reasonably valid and that images in individual cases were properly acquired and evaluated.

The ultimate guide to wisdom about the proper use of neuroscientific evidence is a keen understanding of legal relevance, which in turn requires an equally keen understanding of the legal question at issue. The question in any case, then, is how, precisely, does neuroscience evidence help decide whether an act or mental state criterion was present at the relevant time. No hand-waving about relevance is allowed. For example, a broken brain or a gene-environment interaction that raises the risk of antisocial behavior is *not per se* a mitigating or excusing condition. Such evidence is relevant only if it supports the presence of a genuine excusing or mitigating condition. Whatever rhetorical use an advocate may be able to make of neuroevidence is

distinguishable from whether the evidence is really, as opposed to rhetorically, relevant. The chain of inference from the purely mechanical neurodata to the law's act and mental state criteria must be clear. Unless the neuroevidence can help answer the specific legal question in issue, it is not legally relevant, even if it is scientifically valid. Thus, if there is a disjunct between the subject's behavior and the neuroevidence, actions always speak louder than images, except perhaps in cases of malingering (although the science cannot at present reliably and validly identify malingerers) (this claim is discussed further in the next Part). If the defendant's brain appears broken, but he is a rational agent, he is rational for legal purposes. If the brain appears normal, but the agent is clearly psychotic, the agent is not rational for legal purposes.

For example, Fetal Alcohol Syndrome plays a large role in the Canadian cases (although not in the other samples), but the potentially legally relevant aspects of the disorder are the cognitive and rationality defects, which are behavioral signs, that sufferers demonstrate from an early age. Are the brains of FAS sufferers different from the brains of those without the disorder? Of course. This is just a necessary truth of biological materialism. If the behavior is markedly different, so will be the brain. Brain difference is not *per se* a mitigating or excusing condition, however. If a particular FAS sufferer is somehow sufficiently able rationally to regulate his behavior, then FAS is irrelevant to mitigation or excuse. Moreover, if a FAS sufferer exhibited lifelong cognitive defects, as many do, that sufferer is potentially excusable even if sophisticated neurotechniques cannot identify the brain pathology or brain difference.

Many of the cases in these studies fail to understand the relevance of the neuroevidence. Even if there is clear evidence of brain damage or a neurological disorder, it does not mean that the defendant did not act, lacked *mens rea*, was less culpable, is incompetent, or will be dangerous in the future. All the criteria depend on direct assessment of the offender's behavior. The alleged relevance of neuroevidence to competence determinations, which occurs in many of the samples, is instructive but bewildering. Criminal competencies are behaviorally functional and again defined entirely in terms of mental states. Does the defendant understand the nature of the charges, can he rationally assist counsel, does he understand the consequence of a guilty plea, does he understand the nature of the penalty about to be imposed on him and why it is being imposed? These normative, mental criteria must all be evaluated behaviorally. Either the defendant can perform these tasks to the requisite degree or he cannot.

These are continuum capacities, however, and it may be asked whether neuroscience can help with the gray area, indeterminate cases. The answer is, no, for various reasons. Any brain condition will have heterogeneous consequences. Some people with very broken brains have essentially normal mental functioning. But, cannot group data about people with this condition help us draw inferences at the margin. Once again, the answer is, no, in the present state of neuroscience as a result of the “clear cut” problem already discussed in Part III. In behaviorally unclear cases in which the law needs help the most, neuroscience is least able to furnish it.

A critical reader will be repeatedly struck by how many of the expanded cases either used irrelevant or weak (or non-existent) neuroscience--e.g., to assess competence or whether a defendant suffered from a mental illness--or could have been fully resolved with more careful behavioral evaluation. Of course there can be conflict about the behavioral evidence, but because act and mental state questions must be resolved, it is the behavioral evidence that is doing the real work. And for the reasons given, neuroevidence will seldom be helpful in resolving the gray area cases in which most help is needed. Much is at stake in criminal cases and of course judges would like scientific help to resolve the vexing normative issues they must resolve, but, at present, turning to the neuroscience will do nothing more in most cases than to provide a rationalization for a result the judge wishes to reach on other grounds or to avoid responsibility for having to make the hard decision directly by relying on the expert. Convergent behavioral and neurodata might help solve some of these problems that cannot be resolved with either type of evidence alone, but such convergent line of legally-relevant evidence are very rare.

If a proper framework for the relevance of neuroscience to law is established and if a cautious approach to the science is adopted, I think neuroscience can potentially help refine legal mental state categories, such as *mens rea* and mental disorder through a conceptual-empirical equilibrium in which legal categories guide neuroscientific investigation that in turn then help clarify the legal categories. Neuroscience might also help the fairness and efficiency of criminal law decision making by increasing predictive accuracy. The criminal law already uses predictions for purposes of diversion, sentencing, parole, and the quasi-criminal commitment of some sexual offenders. We have already decided as a normative matter that predictions are acceptable. If neural variables make this practice more accurate at reasonably acceptable cost, that is an advance. Finally, in tandem with behavioral science, neuroscience might help us more accurately understand legally-relevant human capacities, such as the capacity for rationality and for self-control, which would again improve legal policy, doctrine and

adjudication. But all such optimistic outcomes will depend on precise understanding of legal relevance and valid science.

VII. The Ethics of Caution

The new neuroscience is enormously exciting. Investigators are making important discoveries and we appear to be on the threshold of understanding some of the basic mechanisms of the brain, the key to ourselves and our behavior. Consequently, and understandably, many people make exaggerated claims about how much we know and about the relevance and implications for moral, political and legal analysis. Moreover, advocates propose using neuroscientific techniques in situations in which there is substantial and even overwhelming reason to believe the use is not yet valid or otherwise warranted. Many such people suffer from a “syndrome” identified as “brain overclaim syndrome (Morse, 2006).” Neuroscience seems like such powerful, rigorous science that it might wield more influence than it should. Such excess should be avoided to prevent misunderstanding and misuse, which have dangers of their own. I shall give a number of examples to present the problem. The solution to such problems is obvious: caution and modesty in making claims for the implications of the data and technology.

An op-ed in the New York Times purported to demonstrate that brain imaging could teach us a great deal about the “real roots” of political judgment (Iacoboni, et al., 2007). The research base had not yet been peer-reviewed, which raises ethical questions in itself, but the authors also claimed far more understanding of the relation between brain and political judgment and behavior than science possesses. Moreover, it was implicitly suggested that political judgments are simply reducible to the brain and do not have independent validity that can supported by reason. For another example, a psychiatrist speculated on the potential brain abnormalities of public figures, such as President Clinton, and suggested that perhaps all presidential candidates should undergo brain scans (Amen, 2007). Not only is it unlikely that the writer had personally evaluated the subjects of his speculation, but there is not a shred of evidence either that most of the subjects suffered from the alleged abnormalities or that the subjects’ undesirable behavior was the result of these abnormalities.

The US Supreme Court decided from 2005 to 2012 a trilogy of cases concerning the just punishment of adolescents that mandated differential treatment for many serious adolescent offenders (*Roper v. Simmons*, 543 U.S. 551 (2005); *Graham v. Florida*, 2010; *Miller v. Alabama*, 2012). Prior to the decisions, numerous professional organizations, including the American Psychiatric Association (*Roper v. Simmons*, 2005), urged the Court to hold that adolescents were not fully responsible in part because neuroimaging

studies demonstrate that adolescent cortical neurons are not fully myelinated. In the latter two cases, the Court cited the neuroscience rather vaguely and the science was not necessary to the Court's rationale because the Court adopted the controlling rationale in the first case without citing neuroscience. The science was valid, but, with respect, the claims for relevance were not. As we have seen, the capacity for rationality is the essential, behavioral criterion for responsibility. This must be assessed behaviorally and cannot be "read off" from any brain measurement unless the brain variable is precisely correlated with the behavioral criterion in question that has been normatively identified. But we know that there is vast variability in the brain/behavior link and such correlations are well beyond present knowledge and may never be possible as a result of such variation. If agreement existed about the normative behavioral standard and precise correlates not subject to the clear cut problem were discovered, then neuroscience might help resolve close cases behaviorally, but once again, this is far beyond our present capabilities. Moreover, based on common sense and on excellent behavioral science studies that the Supreme Court ultimately cited, we already knew without question that adolescents as a class are on average less rational than adults and that such lesser rational capacity could provide a moral and legal basis for holding them less responsible. It would be best to individuate such decisions, and the Court did so in *Miller*, but not in the other two in which they treated adolescents as a class. Perhaps in the future neuroscientific evidence might help individuate, but yet again, this is far beyond our present abilities. At most, the myelination evidence offered nothing more than an additive, partial causal story about why later adolescents might be less rational on average than adults. The biological difference was *per se* not relevant to the legal criteria.

Whether adolescents are sufficiently less rational on average than adults, to exclude them categorically from the death penalty or from any other punishment is a normative legal question and not a scientific or psychological question. The neuroscience evidence in no way independently confirms that adolescents are less responsible. If the behavioral differences between adolescents and adults were slight, it would not matter if their brains were quite different. Similarly, if the behavioral differences were sufficient for moral and constitutional differential treatment, then it would not matter if the brains were essentially indistinguishable. If the brains were indistinguishable, the most sensible inference would be that neuroscience is not yet sensitive enough to track the behavioral differences, not that we are mistaken about whether behavioral differences exist. If the system were individuating responsibility differences and neuroscience was sufficiently advanced, then the science might help resolve close cases. But as suggested above and below, actions speak

louder than images. If there is a disjunct, except in cases of malingering, we must believe the behavioral evidence and neuroscience cannot identify subjects who are not candid about their mental capacities.

A related, usually unjustified application of neuroscience is the use of imaging evidence to aid criminal justice decision making. For example, simply finding a brain abnormality of any sort does not entail any legal conclusion about responsibility. Partial brain causation is not the equivalent of compulsion. And again, responsibility criteria are behavioral and must ultimately be assessed behaviorally. If there is disjunction between the imaging evidence and the behavior, the behavioral evidence must almost always dominate. At most, the imaging evidence may help us resolve cases in which the behavioral evidence is in dispute, but only if the imaging evidence is relevant to the behavioral criterion in question. At present, that will seldom be true. In the future, however, as suggested above, neuroscientific evidence might help resolve whether and to what degree a subject suffers from a psychotic disorder that bears on responsibility.

The final example of potential misuse is the use of neural lie detection (Greely and Illes, 2007). The use of this technology has enormous criminal justice and civil liberties implications. Private companies have started marketing neural lie detection technology. Limited studies under limited laboratory conditions have indicated some success in detecting intentional misstatements under conditions when nothing is really at stake, but is there sufficient scientific justification for bringing this technology into the public domain? Most informed observers think that neural lie detection has not yet been sufficiently validated for these purposes, and its unregulated use thus has the potential for enormous mischief if people credulously believe that it is more accurate than it actually is. Indeed, in one widely-noted case, a federal magistrate judge excluded neural lie detection evidence in a criminal case on the ground that this method did not meet the legal standard for the admissibility of scientific or technical evidence (U.S. v. Semrau, 2010).

How should the law respond when valid and relevant neuroevidence is inconsistent with the defendant's behavior? Recall that the criminal law's criteria are all behavioral—actions and mental states. Therefore, cases of malingering aside, actions speak louder than images. This is a truism for all criminal responsibility. If the finding of any test or measurement of behavior is contradicted by actual behavioral evidence, then we must believe the real world behavioral evidence because it is more direct and probative of the law's behavioral criteria. For example, if the person behaves rationally in a wide variety of circumstances, the agent is rational even if the brain appears structurally or functionally abnormal. We confidently knew that some people

were behaviorally abnormal, such as being psychotic (grossly out of touch with reality), long before there were any psychological or neurological tests for such abnormalities.

An analogy from physical medicine may be instructive. Suppose someone complains about back pain, a subjective symptom, and the question is whether the subject actually does have back pain. We know that many people with abnormal spines do not experience back pain, and many people who complain of back pain have normal spines. If the person is claiming a disability and the spine looks dreadful, evidence that the person regularly exercises on a trampoline without difficulty indicates that there is no disability caused by back pain. If there is reason to suspect malingering, however, and there is not clear behavioral evidence of lack of pain, then a completely normal spine might be of use in deciding whether the claimant is malingering. Unless the correlation between the image and the legally relevant behavior is very powerful, however, such evidence will be of limited help. Further, although the neuroscience of pain is making advances (Pustilnik, 2015) neuroscience cannot be used at present to diagnose mental disorder because scanning is insufficiently sensitive for these purposes.

If actions speak louder than images and the clear cut problem (see section III above) exists, however, what room is there for introducing neuroevidence in legal cases? Let us begin with cases in which the behavioral evidence is clear and permits an equally clear inference about the defendant's mental state. For example, lay people may not know the technical term to apply to people who are manifestly out of touch with reality, but they will readily recognize this unfortunate condition. No further tests of any sort will be necessary to prove that the subject suffers from seriously impaired rationality. In such cases, neuroevidence will be at most convergent and increase our confidence in what we already had confidently concluded. Determining if it is worth collecting the neuroevidence will depend on whether the cost-benefit analysis justifies obtaining convergent evidence.

For another example, suppose that in an insanity defense case the question is whether the defendant suffers from a major mental disorder such as schizophrenia. In extreme cases, the behavior will be clear, and no neurodata will be necessary. Investigators have discovered various small, but statistically significant, differences in neural structure or function between people who are clearly suffering from schizophrenia and those who are not. Let us assume the validity of these findings although there is reason to be very cautious (Button et al., 2013; Ioannidis, 2011). Nonetheless in a behaviorally unclear case, the overlap between data on the brains of people with schizophrenia and people without the disorder is so great that a scan is insufficiently sensitive to be used

for diagnostic purposes. In short, at present in those cases in which the neuroscience would be most helpful, it has little to contribute. Again, this situation may change if neural markers become more diagnostically sensitive for legally relevant criteria.

VIII. Conclusion: The Need for New Ethical Resources

At present, we have no idea how the brain produces consciousness and enables the mind and action. What if neuroscience (or any other) unlocks those mysteries. This will cause a profound revolution in our understanding of ourselves and may make possible extraordinary interventions in the lives of people, ranging from genuinely reading minds to mind control. Current ethical and legal theory consider people as we understand them today and there is no radical shift yet in our understanding of the person despite the astonishing advances in neuroscience and other sciences. Thus, current theory seems adequate to consider the issues that new technologies produce. If a profound revolution in our understanding of ourselves and biological processes occurs, however, there is no guarantee that current theories will be sufficient to help consider and resolve new quandaries. Let us hope that if this scenario should ever arise, new ethical and legal theory will be adequate to the task.

References

- Adolphs, R. (2015). "The Unsolved Problems of Neuroscience." *Trends in Cognitive Sciences* 19(4): 173-175.
- Aharoni, E., et al. (2013). "Neuroprediction of Future Rearrest." *Proceedings of the National Academy of Sciences* 110(15): 6223-6228.
- Amen, D.G. (2007). "Getting Inside their Heads...Really Inside." *Los Angeles Times*, December 5, A:31.
- Eklund, A., Nichols, T.E. and Knutsson, H. (2016). "Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates." *Proceedings of the National Academy of Science* 113: 7900-7905.
- Barkley, R.A., and Murphy, K.R. (2010). "Impairment in Occupational Functioning and Adult ADHD: The Predictive Utility of Executive Function (EF) Ratings versus EF Tests." *Archives of Clinical Neuropsychology* 25(3): 157-173.
- Bennett, C.M., Wolford, G.L., and Miller, M.B. (2009). "The Principled Control of False Positives in Neuroimaging." *Social Cognitive and Affective Neuroscience* 4(4): 417-422.

Bennett, C.M., et al. (2009). "Neural Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument for Proper Multiple Comparisons Correction." *Journal of Serendipitous and Unexpected Results* 1(1): 1. <http://perma.cc/VU7B-K5DJ>.

Berker, S. (2009). "The Normative Insignificance of Neuroscience." *Philosophy & Public Affairs* 37(4): 293-329.

Berg, J.W., Appelbaum, P.S., Lidz, C.W., and Parker, L.S. (2001). *Informed Consent: Legal Theory and Clinical Practice*, 2nd ed. (New York, Oxford University Press).

Berman, M.N. (2008). "Punishment and Justification." *Ethics* 118(2): 258-290.

Buchanan, A. (2011). *Better than Human: The Promise and Perils of Enhancing Ourselves* (New York: Oxford University Press).

Buchanan, A., Brock, D.W., Daniels, N., and Wikler, D. (2000). *From Chance to Choice: Genetics and Justice* (Cambridge: Cambridge University Press).

Buckholz, J.W., et al. (2015). "From Blame to Punishment: Disrupting Prefrontal Cortex Activity Reveals Norm Enforcement Mechanisms." *Neuron* 87: 1369-1380.

Button, K.S., et al. (2013). "Power Failure: Why Small Sample Size Undermines the Reliability of Neuroscience." *Nature Reviews Neuroscience* 14: 365-376.

Cashmore, A.R. (2010). "The Lucretian swerve: The biological basis of human behavior and the criminal justice system." *PNAS* 107(10): 4499-4504.

Catley, P. and Claydon, L. (2015) "The use of neuroscientific evidence in the courtroom by those accused of criminal offenses in England and Wales." *Journal of Law and Biosciences*. 2:510-549.

Chandler, J. (2015). "The use of neuroscientific evidence in Canadian criminal proceedings." *Journal of Law and Biosciences* 2: 550-579.

Chatterjee, A., and Farah, M. (2013). *Neuroethics in Practice* (Oxford: Oxford University Press).

Chin, J.M. (2014). "Psychological Science's Replicability Crisis and What It Means for Science in the Courtroom." *Psychology, Public Policy, and Law* 20(3): 225-238.

de Kogel, C.H. and Westgeest, E.J.M.C/ (2015) "Neuroscientific and behavioral genetic information in criminal cases in the Netherlands," *Journal of Law and Biosciences* 2: 580-605.

- Faigman, D.L., Monahan, J., and Slobogin, C. (2014). "Group to Individual (G2i) Inference in Scientific Expert Testimony." *The University of Chicago Law Review* 81(2): 417-480.
- Farah, M., ed. (2010). *Neuroethics* (Cambridge, Mass.: MIT Press).
- Farahany, N.A. (2012). "Incriminating Thoughts." *Stanford Law Review* 64:351-408.
- Farahany, N.A. (2015). "Neuroscience and Behavioral Genetics in U.S. Criminal Law: An Empirical Analysis." *Journal of Law and Biosciences* 2:485-509.
- Fins, J.J., Illes J., Bernat J.L., Hirsch. J., Laureys, S., and Murphy, E. (2008). "Neuroimaging and Disorders of Consciousness:Envisioning an Ethical Research Agenda" *The American Journal of Bioethics*, 8:9, 3-12, DOI:10.1080/15265160802318113.
- Fodor, J.A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (Cambridge, Mass.: MIT Press).
- Frances, A. (2009). "Whither DSM-V?" *British Journal of Psychiatry* 195(5): 391-392.
- Garland, B., ed. (2004). *Neuroscience and the Law: Brain, Mind and the Scales of Justice* (New York: Dana Press).
- Garreau, J. (2005). *Radical Evolution: The Promise and Peril of Enhancing our Minds, our Bodies—and what it means to be Human* (New York: Doubleday).
- Ginther M., Bonnie, R. J., Hoffman, M.B. et al (2016). "Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment." *J. Neuroscience* 36(36):9420 –9434.
- Glannon, W. (2007). *Defining Right and Wrong in Brain Science* (New York: Dana Press).
- Glannon, W. (2008). *Brain, Body and Mind: Neuroethics with a Human Face* New York: Oxford University Press).
- Greely, H.T. (2013). "Mind Reading, Neuroscience, and the Law," in S.J. Morse and A.L. Roskies (eds.), *A Primer on Criminal Law and Neuroscience* (New York: Oxford University Press).
- Greely, H.T., and Illes, J. (2007). "Neuroscience-based Lie Detection: The Urgent Need for Regulation." *American Journal of Law and Medicine* 33(2-3): 377-431.
- Greene, J., and Cohen, J. (2006). "For the Law, Neuroscience Changes Nothing and Everything," in S. Zeki and O. Goodenough (eds.), *Law and The Brain* (New York: Oxford University Press).

Harris, J. (2007). *Enhancing Evolution: The Ethical Case for Making Better People* (Princeton: Princeton University Press).

He, B. (ed.) (2013). *Neural Engineering* (New York: Springer Science).

He, B., Gao, S., Yuan, H. and Wolpaw, J.R. (2013). "Brain-Computer Interfaces," in B. He (ed.), *Neural Engineering* (New York: Springer Science).

Hoffman, M. (2014). *The Punisher's Brain: The Evolution of Judge & Jury* (New York: Cambridge University Press).

Holtzheimer, P.E. and Mayberg, H.S. (2011). "Deep Brain Stimulation for Psychiatric Disorders." *Annual Review of Neuroscience* 34: 289-307.

Husak, D., and Murphy, E. (2013). "The Relevance of the Neuroscience of Addiction to the Criminal Law," in S.J. Morse and A.L. Roskies (eds.), *A Primer on Criminal Law and Neuroscience* (New York: Oxford University Press).

Iacoboni, M., et al. (2007). "This is your Brain on Politics." *New York Times*, November 11, 4:14.

Illes, J. (2005). *Neuroethics* (Oxford: Oxford University Press).

Illes, J., and Sahakian, B. (2011). *Oxford Handbook of Neuroethics* (Oxford: Oxford University Press).

Ioannidis, J.P.A. (2011). "Excess Significance Bias in the Literature on Brain Volume Abnormalities." *Archives General Psychiatry* 68(8): 773-780.

Jones, O.D., and Ginther, M. (2015). "Law and Neuroscience," in J.D. Wright (ed.), *International Encyclopedia of the Social & Behavioral Sciences*, 2nd ed. (New York: Elsevier).

Jones, O.D., Schall, J. and Shen, F.X. (2013). *Law and Neuroscience* (New York: Walters Kluwer Law and Business).

Kalivas, P.W., and Volkow, N.D. (2005). "The Neural Basis of Addiction: A Pathology of Motivation and Choice." *American Journal of Psychiatry* 162(8): 1403-1413.

Kamm, F. (2009). "Neuroscience and Moral Reasoning: A Note on Recent Research." *Philosophy & Public Affairs* 37(4): 330-345.

Kane, R. (1998). *The Significance of Free Will* (New York: Oxford University Press).

Kyllo v. U.S., 533 U.S. 27 (2001).

- Levy, N. (2007). *Neuroethics: Challenges for the 21st Century* (Cambridge: Cambridge University Press).
- Libet, B. (1999). "Do We Have Free Will?" *Journal of Consciousness Studies* 6(8-9): 47-57.
- Lieberman, M.D., Berkman, E.T., and Wager, T.D. (2009). "Correlations in Social Neuroscience Aren't Voodoo: A Commentary on Vul et al. (2009)." *Perspectives on Psychological Science* 4(3): 299-307.
- Lulé, D., Noirhomme, Q., Kleih, S.C. et al (2012). "Probing command following in patients with disorders of consciousness using a brain-computer interface." *Clinical Neurophysiology* <http://dx.doi.org/10.1016/j.clinph.2012.04.030>
- Mandavilli, A. (2006). "Actions Speak Louder than Images." *Nature* 444(7120): 664-665.
- Marangell, L.B., Martinez, M., Jurdi, R.A., and Zboyan, H. (2007). "Neurostimulation Therapies in Depression: A Review of New Modalities." *Acta Psychiatrica Scandinavica* 116(3): 174-181.
- McHugh, P.R., and Slavney, P.R. (1998). *The Perspectives of Psychiatry*, 2nd ed. (Baltimore: Johns Hopkins University Press).
- M'Naghtens Case, 10 CL & F. 200, 8 Eng. Rep. 718 (1843).
- Mele, A.R. (2009) *Effective Intentions: The Power of Conscious Will* (New York: Oxford University Press).
- Mele, A.R. (2014) *Free: Why Science Hasn't Disproved Free Will* (New York: Oxford University Press).
- Menninger, K. (1968). *The Crime of Punishment* (New York: The Viking Press).
- Miller, G.A. (2010). "Mistreating Psychology in the Decades of the Brain." *Perspectives on Psychological Science* 5(6): 716-743.
- Moore, M. (2012). "Responsible Choices, Desert-Based Legal Institutions, and the Challenges of Contemporary Neuroscience." *Social Philosophy and Policy* 29(1): 233-279.
- Morse, S.J. (1994). "Culpability and Control." *University of Pennsylvania Law Review* 142(5): 1587-1660.
- Morse, S.J. (2004). "New Neuroscience, Old Problems," in B. Garland (ed.) *Neuroscience and the Law: Brain, Mind and the Scales of Justice* (New York: Dana Press).
- Morse, S.J. (2006). "Brain Overclaim Syndrome: A Diagnostic Note." *Ohio State Journal of Criminal Law* 3(2): 397-412.

Morse, S.J. (2007). "The Non-problem of Free Will in Forensic Psychiatry and Psychology." *Behavioral Sciences and the Law* 25(2): 203-220.

Morse, S.J. (2011). "Lost in Translation? An Essay on Law and Neuroscience," in M. Freeman (ed.) *Law and Neuroscience* 13(28): 529-562.

Morse, S.J. (2012). "New Therapies, Old Problems, or, A Plea for Neuromodesty." *AJOB Neuroscience* 3(1): 60-64.

Morse, S.J. (2013). "Brain Overclaim Redux." *Law and Inequality* 31(2): 509-534.

Morse, S.J. and Roskies, A.L (eds.) (2013). *A Primer on Criminal Law and Neuroscience* (New York: Oxford University Press).

Morse, S.J. (2015). "Neuroprediction: New Technology, Old Problems." *Bioethics Forum* 8(4): 128-129.

Morse, S.J., and Newsome, W.T. (2013) "Criminal Responsibility, Criminal Competence, and Prediction of Criminal Behavior," in S.J. Morse and A.L. Roskies (eds.) *A Primer on Criminal Law and Neuroscience* (New York: Oxford University Press).

Nachev, P., and Hacker, P. (2015). "The Neural Antecedents to Voluntary Action: Response to Commentaries." *Cognitive Neuroscience* 6(4): 180-186.

Open Science Collaboration. (2015). "Psychology: Estimating the Reproducibility of Psychological Science." *Science* 349(6251): aac4716-1-aac4716-8.

Pardini, D.A., Raine, A., Erickson, K., and Loeber, R. (2014). "Lower Amygdala Volume in Men Is Associated with Childhood Aggression, Early Psychopathic Traits, and Future Violence." *Biological Psychiatry* 75(1): 73-80.

Pardo, M.S. & Patterson, D. (2013). *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience* (New York: Oxford University Press).

Parens, E. (2005). "Authenticity and Ambivalence: Toward Understanding the Enhancement Debate." *The Hastings Center Report* 35(3): 34-41.

Presidential Commission for the Study of Bioethical Issues (2015). *Gray Matters: Topics at the Intersection of Neuroscience, Ethics and Society, Volume 2* (Washington, D.C.: Presidential Commission for the Study of Bioethical Issues).

Pustilnik, A.C. (2015). "Imaging Brains, Changing Minds: How Pain Neuroimaging Can Help Transform the Law." *Alabama Law Review* 66(5): 1099-1158.

Rakoff, J.S. (2016) "Neuroscience and the Law: Don't Rush In." *New York Review of Books* LXIII: 30-.

- Rego, M.D.(2016). “Counterpoint: Clinical Neuroscience is Not Ready for Clinical Use” *British Journal of Psychiatry* 208: 312-313
- Roper v. Simmons, 543 U.S. 551 (2005).
- Roskies, A.L (2013) “Brain Imaging Techniques,” in S.J. Morse and A.L. Roskies (eds.) *A Primer on Criminal Law and Neuroscience* (New York: Oxford University Press).
- Roskies, A.L. (2002),”Neuroethics for the New Millenium.” *Neuron* 35(1):21-23.
- Roskies, A.L.(2016). “Neuroethics.” *Stanford Encyclopedia of Philosophy* (Stanford: Stanford University).
- Sandel, M.J. (2007). *The Case against Perfection: Ethics in the Age of Genetic Engineering* (Cambridge, Mass.: Harvard University Press).
- Satel, S. and Lilienfeld, S.O. (2013). *Brainwashed: The Seductive Appeal of Mindless Neuroscience* (New York: Basic Books).
- Schneider. M-H., Fins, J.J. and Wolpaw, J.R. (2012). “Ethical Issues in BCI Research,” in J.R. Wolpaw and E.W. Wolpaw (eds.), *Brain-Computer Interfaces: Principles and Practice* (New York: Oxford University Press).
- Schurger, A., and Uithol, S. (2015). “Nowhere and Everywhere: The Causal Origin of Voluntary Action.” *Review of Philosophy and Psychology* 6(4): 761-778.
- Schurger, A., Sitt, J.D., and Dehaene, S. (2012). “An Accumulator Model for Spontaneous Neural Activity Prior to Self-Initiated Movement.” *Proceedings of the National Academy of Sciences of the U.S.A.* 109(42): E2904-E2913.
- Scott,E. et al (*in press*) “Young Adulthood as a Transitional Legal Category: Science, Social Change and Justice Policy” *Fordham Law Review*
- Sell v. U.S.,539 U.S. 166 (2003).
- Sinnott-Armstrong, W. and Nadel, L. (2010). *Conscious Will and Responsibility: A Tribute to Benjamin Libet* (New York: Oxford University Press).
- Soon, C.S., et al. (2008). “Unconscious Determinants of Free Decisions in the Human Brain.” *Nature Neuroscience* 11: 543-545.
- State v. Randall, 532 N.W.2d 94 (Wis. 1995).
- Strawson, G. (1989). “Consciousness, Free Will, and the Unimportance of Determinism.” *Inquiry* 32(1): 3-27.
- Strawson, P.F. (1982). “Freedom and Resentment,” in G. Watson (ed.), *Free Will* (Oxford: Oxford University Press).

Szucs, B. and Ioannidis, J.P.A. (2016). “Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature.” *bioRxiv* (preprint first posted online Aug. 25, 2016); <http://dx.doi.org/10.1101/071530>

U.S. v. Sell, 123 S. Ct. 2174 (2003).

Vincent, N.A. (ed.) (2013). *Neuroscience and Legal Responsibility* (New York: Oxford University Press).

Vul, E., Harris, C., Winkiekman, P., and Pashler, H. (2009). “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition.” *Perspectives on Psychological Science* 4(3): 274-290.

Wallace, R.J. (1994). *Responsibility and the Moral Sentiments* (Cambridge, MA: Harvard University Press).

Walsh, M.T., and Dinan, T.G. (2001). “Selective Serotonin Reuptake Inhibitors and Violence: A Review of the Available Evidence.” *Acta Psychiatrica Scandinavica* 104(2): 84-91.

Washington v. Harper, 494 U.S.211 (1990).

Wegner, D.M. (2002). *The Illusion of Conscious Will* (Cambridge: MIT Press).

Wittgenstein, L. (1953). *Philosophical Investigations* (New York: The Macmillan Company).

Wolpaw, J.R. and Wolpaw, E.W. (eds.) (2012)., *Brain-Computer Interfaces: Principles and Practice* (New York: Oxford University Press).